

开放人文



Why Beauty Is Truth: A History of Symmetry

[英] 伊恩·斯图尔特 著　王天龙 译

Ian Stewart

对称的历史

上海世纪出版集团



上架建议：数学 科学史

ISBN 978-7-208-09800-8



定价：45.00元

易文网：www.ewen.cc
文景网：www.wenjingbook.com

对称的历史

[英] 伊恩·斯图尔特 著　王天龙 译

世纪出版集团 上海人民出版社

图书在版编目(CIP)数据

对称的历史/(英)斯图尔特(Stewart,I.)著;
王天龙译. —上海:上海人民出版社,2011
(世纪人文系列丛书.开放人文)
书名原文:Why Beauty Is Truth: A History of
Symmetry
ISBN 978—7—208—09800—8

Ⅰ.①对… Ⅱ.①斯… ②王… Ⅲ.①科学家—生平
事迹—世界 Ⅳ.①K816.1

中国版本图书馆CIP数据核字(2011)第017346号

责任编辑 管鸥鹏
装帧设计 陆智昌
美术编辑 肖晋兴



对称的历史
[英]伊恩·斯图尔特 著
王天龙 译

出　　版 世纪出版集团 上海人民出版社
(200001 上海福建中路193号 www.ewen.cc)
出　　品 世纪出版股份有限公司 北京世纪文景文化传播有限责任公司
(100027 北京朝阳区幸福一村甲55号4层)
发　　行 世纪出版股份有限公司发行中心
印　　刷 北京鹏润伟业印刷有限公司
开　　本 635×965毫米 1/16
印　　张 22
插　　页 4
字　　数 238,000
版　　次 2011年8月第1版
印　　次 2011年8月第1次印刷
I S B N 978-7-208-09800-8/G.1417
定　　价 45.00元

出版说明

自中西文明发生碰撞以来，百余年的中国现代文化建设即无可避免地担负起双重使命。梳理和探究西方文明的根源及脉络，已成为我们理解并提升自身要义的借镜，整理和传承中国文明的传统，更是我们实现并弘扬自身价值的根本。此二者的交汇，乃是塑造现代中国之精神品格的必由进路。世纪出版集团倾力编辑世纪人文系列丛书之宗旨亦在于此。

世纪人文系列丛书包涵“世纪文库”、“世纪前沿”、“袖珍经典”、“大学经典”及“开放人文”五个界面，各成系列，相得益彰。

“厘清西方思想脉络，更新中国学术传统”，为“世纪文库”之编辑指针。文库分为中西两大书系。中学书系由清末民初开始，全面整理中国近现代以来的学术著作，以期为今人反思现代中国的社会和精神处境铺建思考的进阶；西学书系旨在从西方文明的整体进程出发，系统译介自古希腊罗马以降的经典文献，借此展现西方思想传统的生发流变过程，从而为我们返回现代中国之核心问题奠定坚实的文本基础。与之呼应，“世纪前沿”着重关注二战以来全球范围内学术思想的重要论题与最新进展，展示各学科领域的新近成果和当代文化思潮演化的各种向度。“袖珍经典”则以相对简约的形式，收录名家大师们在体裁和风格上独具特色的经典作品，阐幽发微，意趣兼得。

遵循现代人文教育和公民教育的理念，秉承“通达民情，化育人心”的中国传统教育精神，“大学经典”依据中西文明传统的知识谱系及其价值内涵，将人类历史上具有人文内涵的经典作品编辑成为大学教育的基础读本，应时代所需，顺时势所趋，为塑造现代中国人的人文素养、公民意识和国家精神倾力尽心。“开放人文”旨在提供全景式的人文阅读平台，从文学、历史、艺术、科学等多个面向调动读者的阅读愉悦，寓学于乐，寓乐于心，为广大读者陶冶心性，培植情操。

“大学之道，在明明德，在新民，在止于至善”（《大学》）。温古知今，止于至善，是人类得以理解生命价值的人文情怀，亦是文明得以传承和发展的精神契机。欲实现中华民族的伟大复兴，必先培育中华民族的文化精神；由此，我们深知现代中国出版人的职责所在，以我之不懈努力，做一代又一代中国人的文化脊梁。

上海世纪出版集团

世纪人文系列丛书编辑委员会

2005年1月

对称的历史

等暮年使这一世代都凋落，
只有你如旧；在另外的一些
忧伤中，你会抚慰后人说：
“美即是真，真即是美”，这就包括
你们所知道，和该知道的一切。

——济慈：《希腊古瓮颂》

查良铮 译

目录

前言

1832年5月13日。一片晨雾中，两个法国人相向而立，手枪响了，这是一个年轻的女人引起的决斗。一个男人在一声枪响后倒在了地上，他受了致命伤。两星期后，他死于腹膜炎，年仅21岁。他被葬在一个不起眼的墓地，一个普普通通的墓穴中。数学和科学史上最重要的理论之一差点同他一起死去。

那位活下来的决斗者已不为人知；而死去的那位，就是艾瓦里斯特·加洛瓦（Évarist Galois），一个政治革命家，一个只写了60页著作的数学迷。但是，加洛瓦留下了一个数学革命的传奇。他发明了描述数学结构中的对称的语言，并推导出其结论。

今天，人们把这种语言叫做“群论”，它被应用于理论数学和应用数学中，并以此支配着自然世界的框架和模式。对称也是前沿物理学的主角，包括极其微小的量子论世界和极其广阔的相对论世界，它甚至给出了一条探寻“万有理论”（Theory of Everything）的途径，一种对现代物理学两大主要分支的数学上的统一。而它只不过开始于一

个简单的代数问题，就是如何根据一些数学线索求解方程式中的“未知数”。

对称不是数字也不是形状，而是一种特殊的变换（transformation），一种移动物体的方式。如果一个物体在经过变换之后看起来与先前相同，那这个变换就是对称。例如，一个正方形在转动 90 度之后看起来与先前就是相同的。

对称理论被扩充之后，成了当今科学解释宇宙及其起源的基本理论。爱因斯坦的相对论的中心原理是，在一切的时空中，物理规律都应该是相同的，即，时空中的运动规律是对称的。量子物理学告诉我们，宇宙万物都是由细小的“基本”粒子构成的。这些粒子的运转受数学方程（自然规律）的支配，而这些规律也具有对称性。粒子以数学的方式变换成完全不同的粒子，这种变换也要遵从这个不变的物理规律。

如果对对称没有深入的数学认识，也就不会有当今前沿物理学那些新近形成的观念。这些认识发端于理论数学，随后便进入了物理学。极其有用的理论可能产生于纯粹抽象的沉思，就像人们经常引用的著名物理学家尤根·维格纳（Eugene Wigner）所说的“数学对自然科学的不可思议的影响”。对于数学，我们经常是取多予少的。

从巴比伦的书记员开始，一直到 21 世纪的物理学家，《对称的历史》向我们讲述了数学家们是如何围绕着对称理论摸爬滚打的；也讲述了那些看似无用的研究如何得出了重要公式，从而打开了观察宇宙的新窗口并革新了科学与数学。此外，关于对称的故事还说明，这些重要理论的文化影响和历史脉络是如何在偶然性的政治与科学巨变中清晰起来的。

*

本书的前半部分看起来与对称理论毫无关系，也很少涉及自然界。这是因为，对称理论并不像人们想像的那样是通过几何学演变成一个优势理论的。实际上，当今物理学与数学中奥妙又不可或缺的对称思想，是从代数中来的。因此本书的很大一部分内容都是讲述对代数方程解法的研究的。这种追索看起来很学究，事实上却非常迷人，我们的很多主角有着不同寻常的人生。虽然数学家经常沉迷于抽象思考，但他们都是人。他们当中的一些人，生活确实被过多地逻辑化了，但我们会不止一次地看到，事实上这些英雄们，都太人性了。我们将看到他们的生与死、爱情与决斗、对优先发明权的激烈争夺、绯闻、醉态和疾病，其间，我们还将看到他们的数学思想如何揭示并改变了我们的世界。

本书从公元前 10 世纪开始，到加洛瓦在 19 世纪将其推向顶峰，追溯了人们一步步征服方程式的过程。这一过程终于在数学家遭遇“五次方”方程时被打断了，被未知的五次幂阻断了。是不是五次方程的某些根本特点让那些方法失效了？还是需要一种类似的，但更有效的方法才能得出其解决公式？是数学家们遇到了真正的障碍还是仅仅因为他们太笨了？

弄懂广为人知的五次方程的解法是非常重要的。问题是，它们是否总能用代数方程式来表述？1821 年，年轻的阿贝尔（Norwegian Niels Henrik Abel）证明了五次方程无法用代数方法求解。他的证明神秘而迂回，而且他只证明了普适解法是不可能的，但没有真正解释为什么。

揭示了五次方程为何无法求解的人是加洛瓦，这种不可能性缘自

方程的对称性。如果方程的对称性通过了加洛瓦检验——就是说，这意味着这个等式是以一种特殊的方式组合起来的，我暂时对这种组合方式不予阐释——这个方程就可以用代数公式求解。如果它们没有通过加洛瓦检验，那么这个公式根本就不存在。

一般的五次方程都不能用公式求解，因为它们的对称是有问题的。

*

这项伟大的发现引出了本书的第二个主题："群"——数学上的"对称微积分"。加洛瓦把代数这一古老的数学传统改造成了研究对称的工具。

在本书的这一部分中，"群"这种词还是让人莫名其妙的行话，这些词的词义在叙述中变得重要时，我会对其进行解释的。但我一般只会用一些简单的术语，这就足以弄明白那一大堆林林总总的条目了。如果你遇到了我没有立即论述的行话之类的东西，那就说明它就只不过是个符号而已，其实际意义并不太重要；有时，这些意义会在你的阅读过程中以某种方式呈现出来。"群"是个关键概念，但其含义也是到了本书的中间部分才出现的。

本书还要讲到数学中一些神奇数字的独特意义。我没有借用物理学的内容，而借用了数学中的 π（希腊字母，pi）之类的数字。比如光速，原则上说，它应该是不定的，但它在我们的宇宙中却偶然地成了186000英里/秒。而 π 的值一直稍稍大于3.14159，这又是永远不变的。

五次方程的无法求解告诉我们，5和 π 一样特殊，它是使对称群集无法通过加洛瓦检验的最小数字。另一个奇特的例子是关于1，2，4，8组成的数列。数学家们提出了一系列相对于复数的常规实数概念

的伸延集，然后是被称为四元数和八元数的东西。这些数字分别由实数的 2 倍、4 倍和 8 倍构成。接下来呢？人们会很自然地想到 16，而事实上，在这个数字系统已经没有其他合理的伸延集了。这个事实非同寻常而且相当深刻。它告诉我们，数字 8 是特殊的。这种特殊性不是表面意义上的，而在于数学本身的潜在结构。

除了 5 和 8 之外，对其他一些数字的论述也使本书显得别具一格，最值得注意的包括 14，52，78，133 和 248。这些奇特的数字是五个"例外李群"（exceptional Lie group）的维，它们的影响遍及整个数学和很大一部分的数学物理学。这些数字在数学的舞台上扮演着关键角色，而一些看起来与它们相差无几的数字，却只不过是些小角色而已。

在 19 世纪末现代抽象代数学形成时，数学家们仅仅是发现这些数字非常特殊。这些数字自身并不重要，重要的是它们在代数基本法则中发挥的作用。与这每一个数字都有关联的，是一个属性独特非凡的，被称做"李群"（Lie group）的数学对象。这些群在现代物理学中起着基础性作用，并且看起来与空间、时间和物质的深层结构都有着联系。

*

这就引出了我们的最后一个主题：基础物理学。物理学家们一直为空间具有三个维度而时间却只有一个维度感到困惑，为什么我们生活于四维时空呢？超弦理论是物理学家们试图将整个物理学统一在一个一致的规律下的最新尝试，物理学家们想知道时空结构是否存在着某种"隐匿的"维度。这种想法看起来很荒唐，但却有着很多的历史先例。多维理论可能是超弦理论中最容易为人们所接受的一面。

超弦理论更具争议性的一面是，它认为仅靠相对论和量子论这两大现代物理学的支柱就可以建构一种新的时空理论。统一这两种相互矛盾的理论将会是一个数学课题，而不是探求新的革命性实验的过程。数学美感被看成是物理学真理的首要条件，这可能是个危险的假设。我们不能忽视物理学领域，因为任何在当今人类的深思熟虑后最终诞生的理论，无论其数学渊源有多深，都离不开与实验和观察的比照。

尽管如此，我们还是有很好的理由走上数学的道路。其一，在令人信服的成熟理论建构起来以前，没人知道该进行什么样的试验；另外，数学中的对称理论在相对论和量子论中都扮演着至关重要的角色，而后两者又缺乏共同点，所以我们必须重视我们能够发现的，哪怕是一点点共同点。空间、时间和物质的可能结构是由对称决定的，而且，一些最重要的可能性似乎也与数学中的独特结构相关。也许，时空结构的性质中，只有为数不多的几种得到了数学的认定。这样，进行数学研究也就顺理成章了。

为什么宇宙看起来这么具有数学性呢？人们提供了很多的答案，但我发现每种答案都不太令人信服。数学思想与物理世界的对称，美感与最重要的数学形式中的对称，是一个深刻然而也许无法猜透的秘密。没人能够说清为什么美即是真，真即是美，我们能做的，不过是对这种关系的无限复杂性的沉思而已。

巴比伦书记员

横贯当今被称为“伊拉克”的那片地域，奔腾着世界上最著名的两条河流，一个引人瞩目的文明就由之兴起。这两条河流发源于土耳其北部，流经数百英里的肥沃平原，汇入了流向波斯湾的惟一河道。在西南，它毗邻阿拉伯高原的干旱沙漠地带；在东北，是崇峻的前托罗斯山脉（Anti-Taurus Mountains）和札格洛斯山脉（Zagros Mountains）。这两条河流就是幼发拉底河与底格里斯河，四千年前，它们的流经路线与今天相差无几，贯穿了亚述（Assyria）、阿卡德（Akkad）以及苏美尔（Sumer）的古老土地。

对人类学家来说，幼发拉底河与底格里斯河之间的地域就是著名的美索不达米亚（Mesopotamias），在希腊语中叫“两河流域”。这一地域总是被恰当地称做文明的摇篮。河流给平原带来了水源，水源使平原变得肥沃，丰富的植被引来了羊群和鹿群，继而引来了食肉动物，人类就在其间狩猎。美索不达米亚平原就是狩猎族群的伊甸园、游牧部落的聚集地。

这里如此肥沃，以至于族群狩猎的生活方式渐渐荒废，被更加有效的觅食策略取代了。公元前 9000 年前后，在新月沃地靠北一些的山冈上，诞生了一项革命性的技术——农业。人类社会的两种基本变革紧随而至：为照看庄稼而定居一处以及维持大量人口的可能性。这一变革促使了城市的发明，在美索不达米亚，我们仍能发现世界上最早的城市国家的人类学遗迹：尼尼微（Nineveh）、尼姆鲁德（Nimrud）、尼普尔（Nippur）、乌鲁克（Uruk）、拉旮什（Lagash）、埃利都（Eridu）、乌尔城（Ur），以及最重要的，空中花园和巴别塔的所在地——巴比伦。4000 年前，农业革命使有组织的社会及其相关标志——政体、官僚机构及军事力量——变得可能。公元前 2000 年到公元前 500 年之间，幼发拉底河畔的这片繁荣景象被笼统地冠以“巴比伦文明”之名，但是广义的“巴比伦文明”还包括苏美尔文化和阿卡德文化。提起巴比伦，第一个著名的发现其实是公元前 2250 年前后阿卡德的萨尔冈（Sargon）的一块黏土书写板，但是，也许那时距离巴比伦人的起源已经有两三千年了。

我们对“文明”一词的起源所知甚少，只知道这是一个用以表述人类组织进入稳定社会的书面语。不过，我们今天的世界仿佛在诸多方面都要归功于古老的巴比伦。尤其是，他们是了不起的天文学家，黄道十二星座、以 360 度为一个循环，以及我们以 60 秒为一分钟、60 分钟为一小时，这些都是他们的发明。巴比伦人需要这些计量方法去实践他们的天文学，因此不得不成为天文学忠贞不渝的女仆——数学的专家。

和我们一样，他们也是在学校里学习数学的。

*

“今天上什么课？”那布（Nabu）问道，把自己包着的午餐放在座位一边。他妈妈总要确保他有足够的面包和肉（一般是羊肉）吃，有时还给他放一块奶酪以丰富食物种类。

“数学，”他的朋友加迈什垂头丧气地回答道，“为什么不是法学呢？法学我学得会。”

数学很拿手的那布从来都不太理解为什么他的一帮同学觉得数学很难。“加迈什，你不觉得法律很枯燥吗？照搬一大堆法律短语，还得把它们背熟。”

毅力坚韧又有个好记性的加迈什笑了。“不，这很简单，你不用非思考不可。”

“这正是我觉得它枯燥的地方啊，”他的这位朋友说道，“可是数学——”

“——太难了。”希姆巴巴加入了争论，他刚刚来到“泥板学院”，一如往常地迟到了。“那布，我是说，我该怎么做这道题啊？”他指着泥板上的一道家庭作业习题。“将一个数乘以它自己，然后再加上这个数的2倍，结果是24，这个数是几？”

“是4。”那布回答说。

“真的吗？”加迈什问道。希姆巴巴说：“是的，这我知道啊，但是你是怎么得出来的呢？”

那布艰难地引导着他的两位朋友完成了他们的数学老师上周教给他们的演算过程。“2的一半加上24，等于25，然后开平方，得5。”

加迈什困惑地举起双手说：“我从来都没有真正学会过开平方，那布。”

“啊哈！”那布叫道，“我们就快明白了！”在他的两个朋友看来，他有点发疯。“加迈什，你的问题不是解方程，而是开平方！”

“都有问题。”加迈什咕哝着。

“但首先是开平方。你必须一步一步地学，就像泥板学院的长者一直告诉我们的那样。”

“他总是告诉我们别弄脏了衣服，”希姆巴巴反驳道，“但我们一点也没有注意到——”

“这不一样。这是——”

“这是很糟糕的！”加迈什悲叹道，“我怎么也成不了一个书记员的，我爸爸会把我揍到屁股坐不下去为止，我妈妈则会用乞求的眼光看着我，让我为了家庭努力学习。但是我就是学不会数学！我能记住法典，它很有意思。比如说‘如果一位先生的妻子为了另一个男人杀害了她的丈夫，那么他们应该把她钉在木桩上’。这才是我说的值得一学的东西，而不是那些枯燥的开平方。”他平息了自己的呼吸，但手还因情绪激动而颤抖着，“方程式、数字——我们惹这些麻烦干什么？”

“因为它们有用，”希姆巴巴回答道，“记住那些割奴隶耳朵的法律玩意儿又有什么用？”

“有用！”加迈什说，“惩罚他打人。”

“伤害了一个平民的眼睛，”希姆巴巴怂恿道，“你要赔偿他——”

“一个银迈纳（mina）。”加迈什说。

“打断了一个奴隶的骨头呢？”

“赔偿主人这个奴隶的价钱的一半。”

希姆巴巴开始发难：“那么，如果这个奴隶值60谢克尔（shekel），你就必须算出60谢克尔的一半是多少。如果你想执行法律，就得会数学。”

“是30。”加迈什立即回答道。

“看看，”那布大喊道，“你会数学啊！”

“我才不需要数学呢，这很明显。”未来的律师挥赶着空气，借以发泄自己激烈的情感，“如果这是关于现实问题，这样的数学我做，但不是矫揉造作的开平方。”

“你测量土地时就会用到开平方。”希姆巴巴说。

“不错，但是我上学并不是想成为一个税务员，我爸爸想让我当一个书记员，就像他一样，”加迈什强调说，“我不明白为什么各种数学都要学。”

“因为它有用。”希姆巴巴重复道。

“我觉得这不是真正原因，”那布平静地说，“我想它关乎真与美，关乎得到答案并确知它的正确性。”但是朋友们的表情告诉他，他们并不信服。

“对我来说它关乎得到一个答案并且知道它是错误的。”加迈什叹气道。

“数学之所以重要是因为它既真且美，”那布坚持道，“开平方是解方程的基础，它们可能没太大用，但这并不重要。它们有自身的重要性。”

加迈什正要说出不合宜的话来，可他发现老师走进了教室，就用一声咳嗽掩盖了自己的窘态。

“早上好，孩子们。”老师声音洪亮地说。

“早上好，老师。”

“让我来看看你们的家庭作业。”

加迈什叹着气，希姆巴巴紧张兮兮，那布却依然不动声色，这是个好办法。

*

且不论这场对话完全是虚构的，也许这其中最惊人的东西恰恰被我们遗漏了，那就是这场对话发生在公元前1100年前后的巴比伦这座神奇的城市。

我只是说可能有过这样一场对话，但没有关于这三个名叫那布、加迈什和希姆巴巴的男孩的任何证据，让我们先不管这场对话的记载吧。但是数千年来，人性是一样的，而我这个故事的事实背景也有着岩石般坚实可靠的证据。

我们对巴比伦文明所知甚多，因为他们的记录都是用被称为楔形文字的奇特的楔形符号写在潮湿的黏土板上的。当黏土在巴比伦的阳光下被晒硬，这些铭文几乎就已变成不可毁坏的了。如果存放黏土板的建筑物不小心着火（这种情况时常发生），这些黏土板受热之后就变成了陶器，则会保存得更加长久。

沙漠的最终掩埋会无限期地保留下这些记录。这就是为什么人类历史总要从巴比伦写起。

人类一步步了解对称的故事也是从这里开始的，同时开始的还有它在系统化、量化了的理论中的具体产物——每一部分都像牛顿和莱布尼茨的微积分一样强大的对称微积分。如果有一台时光穿梭机或者一块更古老的泥板，我们无疑会追溯到更早。但是就历史记载向我们所展示的来看，正是巴比伦数学让人类走上了通向对称理论的道路，这对我们如何看待物理世界也有着深远的意义。

*

数学建立在数字上，但并不仅限于数字。巴比伦人拥有更有效的符号系统，与我们使用的“十进制”（decimal，以 10 的幂为基础）不同，他们使用的是“六十进制”（sexagesimal，以 60 的幂为基础）。巴比伦人熟悉直角三角形和类似于我们如今称做毕达哥拉斯定理（Pythagorean theorem）的东西，但是并没有像他们的希腊继承者那样用逻辑证明去支持其来自于经验的发现。他们将数学用于天文学这一更高目的，也可能是用于农业和宗教，以及较普通的商业和税务问题。数学思想不断揭示着自然界的奥秘并辅助着人类事务，它的双重角色像一条金色丝带贯穿着整个数学史。

巴比伦数学家的最重要之处在于，他们开始懂得如何解方程。

方程是一种根据相关条件求未知数值的数学方法，“给出一个未知数的已知条件，求这个未知数”。一个方程就是一个数字的谜，我们并不知道这个数字是什么，只知道一些跟这个数字有关的东西，而我们的任务就是通过求未知数来解开这个谜。这项游戏看起来就像是从几何中的对称观念中分离出来的，但是在数学中，在某种前后关系中取得的发现常常可以开启另一种完全不同的关系。正是这种关联性赋予数学如此强大的理性力量，这就是为什么为商业目的而发明的数字体系也可以向古人揭示出行星和恒星的运动规律的原因。

“一个数的 2 倍是 60，这个数是几？”这道题也许很简单，即使你不是个运算天才，也知道这个数是 30。更难一点的，“一个数乘以它自己，再加上 25，所得结果是这个数的 10 倍，这个数是几？”通过试错，你也可以得知答案是 5，但是试错并不是解方程的有效方法。如果我们把 25 换成 23 或者 26，又该怎么解呢？巴比伦数学家是不

屑于使用试错法的，因为他们已掌握了一个更深刻、更强大的秘密，一种规范、标准的解方程步骤。就我们所知，他们是第一批发现这种技巧的人。

*

巴比伦的部分奥秘出现在大量的圣经引述中。我们都知道发生在尼布甲尼撒国王（King Nebuchadnezzar）统治时期的“狮洞中的丹尼尔”的故事。但是后来，巴比伦几乎成了一种虚构，一座消失了很久的城市，毁坏得无法修复，或许压根就没有存在过。直到大约两百年前，人们差不多都是这么想的。

1881年，克劳迪乌斯·里奇（Claudius Rich）第一次对伊拉克墙基碎石作出科学研究。他在距巴格达向南60英里的幼发拉底河畔进行了勘测，一眼就认定那是巴比伦的遗址，并雇用工人对这片废墟进行了挖掘。他们发现了砖块，写着楔形文字的泥板，通过在黏土上滚动印出文字与图案的漂亮的圆筒印章（cylinder seals），还有堪与达·芬奇和米开朗琪罗比肩的无名艺术家的宏伟雕刻作品。

但最令人感兴趣的还是成堆的写着楔形文字的碎泥板。幸运的是，早期的人类学家们认识到了它们的潜在价值，并将其妥善地保存了起来。上面的笔迹被译读出来之后，这些泥板也就成了关于巴比伦人生活和事务的信息宝库。

泥板以及其他遗物告诉我们，古代美索不达米亚的历史是漫长而复杂的，它牵涉到好几种不同的文化。我们习惯用“巴比伦”这个词把它们全部概括起来，包括那个以巴比伦城为中心的独特文化。但是，美索不达米亚文明的中心是一再迁移的，它就这样选中了巴比

伦，又将它遗弃了。人类学家将巴比伦的历史分为两个主要时期，从公元前2000年到公元前1600年为古巴比伦（Old Babylonian）时期，从公元前625年到公元前539年为新巴比伦时期（Neo-Babylonian）时期。在古巴比伦时期和新巴比伦时期之间是古亚述（Old Assyrian）、喀西特（Kassite）、中亚述（Middle Assyrian）和新亚述（Neo-Assyrian）时期，其间，巴比伦被外来者统治着。此外的500年或更长的一段时间里，巴比伦数学在叙利亚得以延续，这段时期被称为塞琉西（Seleucid）王朝。

巴比伦文明的生命力比它所在的社会更加顽强，它几乎毫无改变地延续了1200年，只是有时在政治动荡的时期发生过暂时的中断。巴比伦文化与其他几种文化相比，有不少独特之处，它的存在可能远早于我们所知的最早记录。尤其是一些数学技巧，其可靠记录可以追溯到公元前600年左右。因此，本章的几个主角——被我命名为那布-沙玛什（Nabu-Shamash）的想像中的书记员以及他的好朋友，我已经在前面简单讲述了他们的初学阶段——被设定生活在公元前1100年前后，生于尼布甲尼撒一世统治时期。

我们故事中的另外几个主角，是真正的历史人物，他们的故事是有明确的文献记录的。从古老的巴比伦保存下来的数以百万计的泥板中，除了皇族和军事将领以外，几乎没有关于个人的文献记录。所以，那布-沙玛什必定是基于我们对巴比伦人日常生活的了解，通过合理的推测描摹出来的。没有任何一种发明可归功于他们，但他们一定接触了巴比伦知识中与对称理论的故事密切相关的各个方面。有很好的证据显示，巴比伦书记员都受过全面的教育，而数学就是其中的重要组成部分。

我们虚构出来的名字是两个真实的巴比伦名字的结合体，即书写

之神那布（Nabu）和太阳神沙玛什（Shamash）。在巴比伦文明中，用神的名字为普通人命名并不稀罕，但是同时使用两个神的名字，在当时也会被认为有点过分。但是由于叙述的需要，我们还是应该把他们叫得特别一点，这比叫他们“书记员”更有感觉。

那布－沙玛什出生时的巴比伦国王是尼布甲尼撒一世，他是第二伊辛王朝（Second Dynasty of Isin）最重要的君主。他不是圣经列王中与他同名的那位国王，那位圣经中经常被提到的国王是尼布甲尼撒二世，是尼布甲尼撒的后裔，其统治是在公元前605年至公元前562年。

尼布甲尼撒二世的统治代表着巴比伦的最高繁荣期，在物质上和疆域上皆然。这座城市在他的同名者统治时期也很繁荣，巴比伦的势力范围扩张到了阿卡德周围以及北部的山区。但是，阿卡德在阿舒尔－列沙－伊什（Ashur-resha-ishi）和提格拉特－帕拉沙尔（Tiglath-Pliser）统治时期摆脱了巴比伦的控制，并通过对环绕其三面的山区和沙漠部落的军事行动强大了起来。所以，那布－沙玛什最初生活在巴比伦的一个稳定时期，但是在他长成一个青年之后，巴比伦之星开始暗淡，生活也就变得动荡不安了。

*

那布－沙玛什生于巴比伦古城一个典型的“上流阶层”家庭，距离利比尔－希加拉运河（Libil-hegalla Canal）不远，并且紧邻伊什塔尔城门（Ishtar Gate），这个城门是由雕刻成公牛、狮子，甚至是龙的奇异形状的五彩砖块装饰的庄严入口。穿过伊什塔尔城门的大道令人惊叹，其宽度达到了20米；它的石灰石路面铺在沥青路基上，地基则是用砖块修建的。这条路的名字叫“不可战胜”（May the enemy

not have victory）——这是一个典型的巴比伦大街的名字，但它通常因“游行大道”（Processional Way）而为人所知，在颂扬马杜克神（Marduk）的仪式命令颁布之后，教士们就通过它进行游行活动。

巴比伦的家庭住宅是用泥砖建造的，遮挡阳光的墙壁有六英尺厚。外墙的开口不多，大体上只有一个临街的出口，房子有三层，由较轻的材料建成，通常是木头，大多用在顶层。各个家庭中都有很多的奴隶，负责家庭的日常事务。他们住在厨房的旁边，就在入口的右侧。家人的卧室则在左侧，有一个狭长的起居室、几间卧室和一个浴室。在那布-沙玛什时代是没有浴缸的，那些保留下来的浴缸是属于其他时代的。因此，那些奴隶们就得往洗浴者的头上和身上浇水，类似于当今的淋浴。这些住宅都有一个露天的中心庭院，背后则是库房。

那布-沙玛什的爸爸是一个国王的法庭里的公务员，这位国王的名字不为人知，其统治要早于尼布甲尼撒一世。爸爸的职责都是些公差：他负责管理一个地区，要维持法律和秩序，主持灌溉农田，收缴必需的税款。他还接受过书记员训练，因为在巴比伦，读写能力和运算能力是一个人为公民服务的基本技能。

学校被称作“泥板学院”，这个名字也许是从那些被用于书写和运算的黏土板来的吧。学校有一位校长，通常被叫做“专家”或者“泥板学院的长者”；有一位班级教师，他的任务是让孩子们规范自己的行为；有教授苏美尔语和数学的专业教师，还有一位被叫做“老大哥”的教官，负责维持秩序。在一个月的30天中，那布-沙玛什有24天住在家里，白天到学校上学。剩下的有三天是娱乐时间，另外三天是宗教节庆。

那布-沙玛什从苏美尔语开始学起，尤其是书写形式的苏美尔语。

他们要学习词汇和语法，还要进行大量的抄写：法律词组、专业术语和姓名。而后他们就开始学习数学，这时，他学习的内容就变成我们故事的中心了。

*

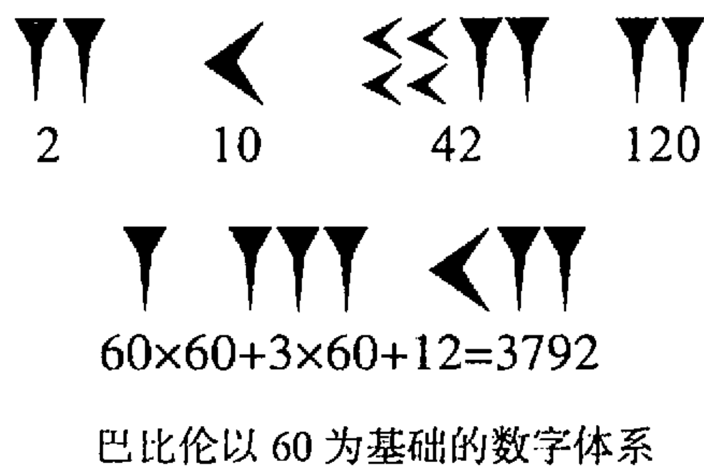
那布－沙玛什为什么要学习呢？除了学究气的哲学家、逻辑学家和专业数学家，一个数（number）对任何人来说都只是一串数字（digit）。我写这本书的时间是2006年，2006就是由四个数字组成的一个串。但是学究们会跳出来提醒我们，这个数字串根本就不是数，而是这个数的符号，是符号的堆砌。我们熟悉的十进制体系仅仅用到了十个数字，即从0到9，用来表示所有的数，不论这个数有多大。经过扩展，这个体系也可以表示非常小的数；更重要的是，它还可以表示高度精确的数量单位。比如光速，最精确的观测数据显示，它接近于186282.397英里/秒。

我们对这些符号太过熟悉了，以至于忘记了这有多么了不起，以及我们第一次遭遇它时想要有所领会是多么困难。所有其他特性都建基于一个关键特性：一个符号的数值，取决于它在其他相关符号中所处的位置。比如2，离开其前后关系，符号2没有任何的固定意义。在表示光速时，紧靠在小数点前面的“2”就表示“二”。但是，这个数中出现在另一处的“2”则表示“两百”。在日期2006中，这个同样的符号则表示“两千”。

如果在我们的字母体系中，每个字母的意义都取决于这个单词中的其他字母，我们就极度郁闷了。设想一下，如果“alphabet”这个单词中的两个a具有完全不同的意义要怎么读。但是，位置记数法

（positional notation）对于数来说方便又有效，所以几乎无法想像有人会采取其他的表示方法。

但并非一直以来都是如此。我们这种表示方法至多有 1500 年的历史，它被引进欧洲也只有 800 年多一点的时间。甚至今天，不同的文化还在用不同的符号表示这相同的十个数字——看看埃及的纸币就明白了。但是，古代文明却用各种不同的方式书写数字。我们最熟悉的是罗马数字体系，其中 2006 变成了 MMVI；在古希腊语中表示为 β̄ς。我们的 2，20，200 和 2000，罗马字体写做 II，XX，CC 和 MM，希腊语中写作 β，κ，σ 和 β̄。



巴比伦以 60 为基础的数字体系

巴比伦文化最早使用了与我们今天类似的位置计数法。但这二者之间有一个本质性区别，在十进制体系中，一个数字每向左移动一位，它表示的数值就要乘以 10。所以，20 就是 2 的 10 倍，200 是 20 的 10 倍。在巴比伦体系中，数字每向左移动一位则要乘以 60。所以，“20”就是 60 的 2 倍（我们记做 120），那么“200”就是 60 的 60 倍的 2 倍（我们记做 7200）。当然，他们并不用“2”这个符号，而是用两个竖长的楔子状的符号表示“二”，如上图所示。从“一”到“九”的数被写做相应的一组竖长的楔子。大于“九”的数，他们用另一种符号来表示，那是一个侧放着的楔子，表示“十”，他们把这些符号组合起来，表示“二十”、“三十”、“四十”和“五十”。因此，我们

的“42”就是四个侧放着的楔子加上后面的两个竖长的楔子。

由于某些不可确知的原因，这个数字体系到了59就停止了。巴比伦人并没有用六个侧放着的楔子表示60，而是掉转头来求助于那个曾被用来表示“一”的竖长楔子，并用其表示“一倍的六十”，而两个这样的楔子就表示120，它们也可能表示“二”。它们到底代表什么还要看其前后关系，要看它们在相关符号中的具体位置。如果先是两个竖长的楔子，中间一个空格，然后又是两个竖长的楔子，那么前面的两个符号表示“一百二十”，而后面的“两个”就分别表示我们的22中的两个符号2的含义“二十”和“二”。

这种表示方法还被延伸用于表示很大的数字。一个竖着的楔子可以表示1，60，或者60×60=3600，或者60×60×60=216000，如此等等。上图最下面的一组楔子表示60×60+3×60+12，而我们会将其写做3792。这里的一个重要问题是，这种标记方法有时会显得含混。如果你看到两个竖长的楔子，它们到底表示的是2，还是60×60×2，还是10×60×60+2×60呢？在亚历山大大帝时期，巴比伦人就用写在缝隙里的一对倾斜的楔子表示数值的不存在，以此克服了这种含混；事实上，他们发明的是表示“零”的符号。

巴比伦人为什么使用六十进制体系而不使用十进制体系呢？也许是60这个数字的有用特征触动了他们吧：它的除数很多。2，3，4，5和6都能将其整除，10，12，15，20和30也能整除它。这种特征在几个人分东西时，比如分谷物和土地时，是非常讨人喜欢的。

最后还有一个决定性的原因——巴比伦人计量时间的方法。看起来他们觉得把一年分为360天很方便，虽然他们都是高超的天文学家，知道一年365天比较准确，$365^1/_4$天则更加准确。360=6×60这种数学关系的诱惑力太大了。但是，应用于时间时，巴比伦人搁置了向左

移动一位就乘以 60 的法则，取而代之以 6，这样，本该是 3600 的就也被表述成了 360。

对 60 和 360 的重视一直延续至今：一个完整的圆是 360 度（一度就是巴比伦人的一天），一分钟有 60 秒，一小时有 60 分钟。古老文化的惯例有着令人难以置信的延续能力。我发现在当今这个电脑书写的壮观时代，电影制作人仍然用罗马数字为他们的作品标注日期。

*

除了“零”的表示方法以外，这些东西那布-沙玛什在受教育的初期应该都学过了。他已经能够熟练地把数以千计的小巧的楔形文字快速地刻写在黏土板上了。就像今天的小学生们吃力地将一个整数化为分数和小数一样，那布-沙玛什也要面对用巴比伦人的方法表示二分之一和三分之一这样的问题，甚至还要面对天文观测数据这一苛刻现实而决定的更复杂的重分（subdivision）问题。

为了避免整个下午都花费在画楔子上，学者们用新方法和旧方法混合表示那些楔形数字。他们把小数写成一组又一组小楔子，把它们用逗号隔开。所以最后的组合就被写成了 1，3，12。这种写法大大节省了费力的排字工作，并且更易识读，所以我们都赞同学者的方法。

巴比伦的书记员到底是怎样写“二分之一”的呢？

在我们的数学中，解决这个问题的方法有两种，要么把它写成分数 1/2，要么借用小数点，写成 0.5。分数表示法更直观，历史也更久；小数表示法比较难于掌握，但这使它更便于计算，因为这种符号体系是从“位值”（place-value）法则延伸出来的。0.5 中的符号 5 表示“5 被 10 等分”，而 0.05 中的 5 表示“5 被 100 等分”。符号向左移

动一位，就是乘以 10；向右移动一位，就是除以 10。

因此，小数运算与整数运算一样，只不过你必须密切注意着小数点的位置。

巴比伦人也有着同样的思想，只不过是以 60 为基础。分数 1/2 就是若干个 1/60。正确的数字很明显是 30/60，所以他们把“二分之一”写成 0;30，学者们用分号表示“六十进位点”（sexagesimal point），而在楔形符号中则是在一个空格之后重新开始。巴比伦人处理过一些相当高级的计算：比如，2 的平方根是 1;24，51，10，其误差不到十万分之一。他们把这种精确性很好地运用到了理论数学和天文学中。

*

那布－沙玛什所学习的技巧中，最激动人心且与本书的主题对称相关的，是二次方程的解法。我们对巴比伦人的解方程方法所知甚多。我们所知的大约 100 万块现存的巴比伦黏土板中，有 500 块是关于数学的。1930 年，东方学家奥托·诺伊格保尔（Otto Neugebauer）发现，其中的一块泥板演示了一个对我们今天称之为二次方程的东西的完整解法。这些方程都包含一个未知量和它的平方，还有很多具体数字。如果没有平方，这个方程就被叫做“线性方程”（linear），而且，这样的方程最容易求解。含有未知数的立方（自身乘以自身，再乘一次这个未知数）的方程叫“三次方程”（cubic）。巴比伦人仿佛掌握了某些三次方程的近似解法，这种解法是以数表（numerical table）为基础的。但是，我们可以肯定的只有这些数表本身。我们只能推测它们的用途，而解三次方程的可能性最大。但是，诺伊格保尔研究过的那些泥板证明，巴比伦的书记员们已经征服了二次方程。

大约4000年前，有一个典型的问题："如果一个正方形的面积减去边长等于14,30，求这个正方形的边长。"这个方程包含着未知数的平方（正方形的面积）和这个数本身。换言之，这就是要读者解二次方程。这块泥板几乎立刻就给出了答案："取1的一半，得0;30，用0;30乘以0; 30，得0;15，加上14,30得14,30；15。这是29;30的平方，现在用29;30加上0;30，结果是30，就是正方形的边长。"

这到底是怎么回事呢？我现在用现代符号把步骤写出来。

- 取1的一半，是0;30。　　1/2
- 0;30乘以0;30，得0;15。　　1/4
- 将得数加上14,30，得14,30;15。　　$870^1/_4$
- 这就是29;30的平方。　　$870^1/_4=(29\ ^1/_2)\times(29^1/_2)$
- 现在用0;30加上29;30。　　$29^1/_2+1/2$
- 得出结果是30，就是正方形的边长。　　30

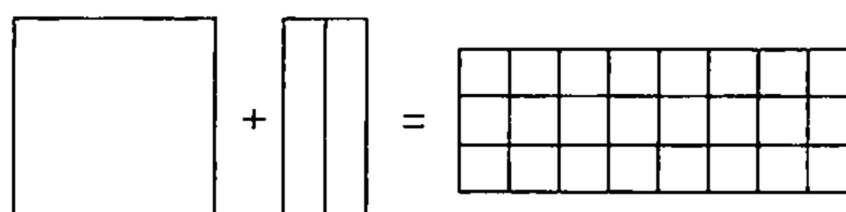
最复杂的步骤是第四步，找到一个其平方等于$870^1/_4$的数（$29\ ^1/_2$）。$870^1/_4$的平方根是$29^1/_2$。开平方是解二次方程的主要工具，当数学家想要用相似的方法解决更复杂的问题时，现代代数就诞生了。

稍后我会用现代代数符号解释这个问题的。但是，我们必须认识到，巴比伦数学家并没有使用这样的代数公式，而是以典型例题的形式描述了一个特殊的步骤，他们就是通过这个步骤得出答案的。但是他们很清楚地知道，虽然数字可以变，同样的步骤却总会有效。

总而言之，他们知道如何解二次方程，他们所用的方法也是我们当今还在使用的，虽然不是通过他们所表达的那种形式。

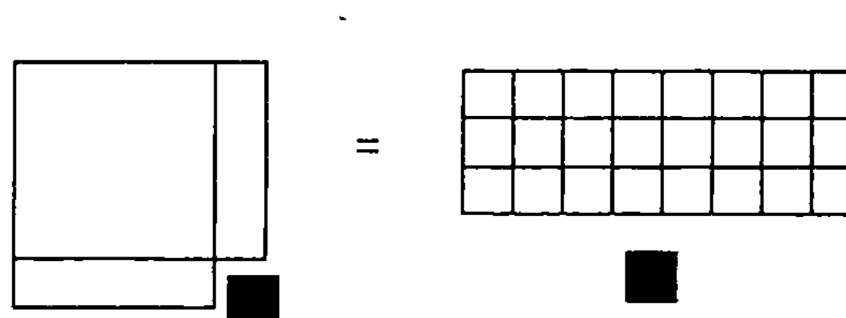
*

巴比伦人是怎样发现这种二次方程的解法的呢？并没有直接的证据表明。但他们好像是通过进行几何式思考来解决的。举个能用这种方法解决的比较简单的问题，假设我们发现一块泥板上这样写着："一个正方形的面积加上它的两条边长，所得之和是24，求正方形的边长。"用较现代的说法就是，未知数的平方加这个数的2倍等于24。我们可以用图来表示这个问题：



一个二次方程的几何示意图

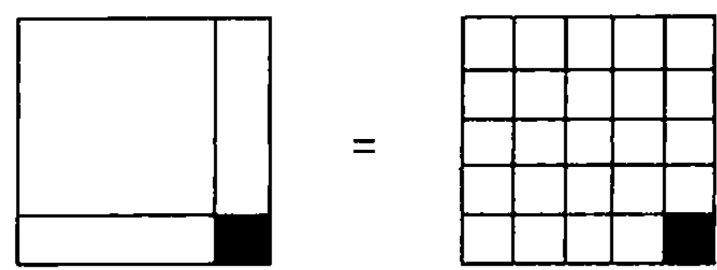
等号左边的正方形和长方形的垂直长度对应着未知数，等号右边的那些小正方形代表单位长度。如果我们把竖着的矩形分为两半，然后把它们粘在正方形上，我们就得到了一个正方形缺一角的图形。图形预示着，我们应该通过在等式的两边同时加上那丢了的一角，"把正方形补充完整"：



补全正方形

这样，等号左边就有了一个正方形，等号右边就有了25个单位

正方形。我们将它们重新排列成一个 5×5 的正方形：



现在，这种解法就很清楚了

因此，未知数加 1，然后平方，结果等于 5 的平方。通过开平方，可知未知数加 1 等于 5，即使你不是运算天才也一样知道未知数是 4。

这种几何描述准确符合了巴比伦人解二次方程的过程。泥板上更复杂的例子用的也是同样的解法。泥板只是陈述了这种方法，而没有说这种方法是怎么来的，但是，这些几何图形给出了间接的证明。

家喻户晓的人

世界上很多最伟大的数学家都生活在埃及的亚历山大城。这座城市起源于尼罗河西岸的五个殷实的绿洲之中。这几个绿洲镶嵌在埃及的西部沙漠里，其中之一就是斯瓦（Siwa），因其冬涨夏落的盐湖而著名。盐败坏了土壤，也给人类学家们带来了很大的麻烦，因为这些盐渗入了古老的岩石和泥砖遗迹之中，慢慢毁坏了建筑物的结构。

斯瓦最著名的旅游景点是阿古米（Aghurmi），一座从前为阿蒙神（god Amun）修建的神庙。阿蒙神太神圣了，以至于他的面貌完全是抽象的，但他与另一个更加具象的实体有关，就是太阳神拉（god Re）。斯瓦的阿蒙神庙建于埃及的第二十六个朝代，它是一个著名的神谕宣示场所，特别是当它与两个重大的历史事件联系起来时。

第一个事件是征服了埃及的波斯国王冈比西斯二世（Cambyses Ⅱ）大军的覆灭。据说这件事发生在公元前523年，冈比西斯二世打算借用阿蒙神的神谕使自己的统治合法化，于是他向西部沙漠中派遣了一支军队。大军到达了巴哈日雅（Bahariya）绿洲，但却在前往斯瓦的

途中被沙暴摧毁了。很多古埃及学学者都怀疑这支“迷失的大军”只是一个神话，但是2000年，阿勒旺大学一个寻找石油的小组，在这一地区发现了衣物、金属和人的遗体，并猜想这就是那支大军的遗迹。

第二个事件是发生在两个世纪后，这可是史实：那就是亚历山大大帝一次致命的到访，他重蹈了冈比西斯二世的覆辙。

*

亚历山大是马其顿国王菲利普二世（King Philip Ⅱ）的儿子。菲利普的女儿，马其顿的克娄巴特拉（Cleopatra of Macedon），嫁给了伊庇鲁斯（Epirus）的亚历山大国王，菲利普就在此过程中被暗杀了。凶手可能是菲利普的同性恋人普萨尼亚斯（Pausanias），他因菲利普国王毫不理睬他的埋怨以及其他举动而心烦意乱。这场谋杀也可能是波斯国王大流士三世（Darius III）的阴谋。如果真是这样的话，那可是适得其反了，因为马其顿军队马上宣布亚历山大为他们的国王，这位20岁的著名统治者继续征服着已知世界的绝大部分地区。在此过程中，他于公元前332年兵不血刃地征服了埃及。

为了巩固其统治，亚历山大带着他法老身份的证明，前往斯瓦朝圣，他想求问神谕自己是不是上帝之子。他独自一人拜见了神使，回来后正式宣布了神使的定论：神谕认定，他的确是上帝之子。这一定论成了他的权威的最高源头。而后，便有谣传称神谕显示他是宙斯之子。

我们并不清楚埃及人在这种这种薄弱的证据面前是不是真的信服了。也许是他们发现，在亚历山大强大的军队的控制下，赞同他所编造的故事还是比较明智的；也可能是埃及人早就厌恶了波斯人的统治，因此转而去欢迎罪恶较小的亚历山大。他受到埃及的前首都孟菲

斯人民的热烈欢迎正是由于这个原因。无论历史背后的真相如何，从那时起，埃及人就崇敬地称亚历山大为国王。

在去斯瓦的路上，亚历山大被地中海与后来被称为玛瑞提斯（Mareotis）的湖泊之间的乡野迷住了，他决定在此建一座城市。他庄重地把这座城市命名为亚历山大（Alexandria）。它是由希腊建筑师多诺克拉提斯（Donocrates）设计的，而其基本蓝图却是亚历山大自己绘制的。一些人将这座城市诞生的时间定为公元前 331 年 4 月 7 日；也有人对此提出过质疑，他们认为城市诞生的时间应该是接近公元前 334 年。亚历山大没来得及看到自己的作品，第二次去的时候，就被埋葬在了那里。

虽然这些被时间镀金的传奇一直流传着，但事实真相却可能更为复杂。现在看来，那座后来被称为亚历山大的城市在亚历山大到来之前就已经存在了。古埃及学学者很早就发现，有很多的碑文并不是完全可信的。比如，拉美西斯二世（Ramses Ⅱ）把卡纳克（Kanark）地区的神庙用涡纹装饰得富丽堂皇，但是这些神庙的很大一部分是他的父亲塞提一世（Seti I）建造的，但是，父亲的事迹（并不总是无关紧要的）却只能在儿子的碑文中找到。这种篡夺很常见，而且常常被认为是无可厚非的。相反，让祖先的雕像受辱，比如划伤法老雕像的脸，通过抹去祖先至高无上的身份来剥夺他死后的地位，才是最为人不齿的。

亚历山大在古老的亚历山大的建筑上刻满了自己的名字。也可以说，他把自己的名字刻进了这座城市。那里的法老们只是篡夺了某一座建筑或者纪念碑，而亚历山大则篡夺了整座城市。

亚历山大变成了一个重要港口，连接尼罗河的支流与一条流入红海的运河，并由此而与印度洋和远东地区相连。它变成了一个知识中心，拥有一座著名的图书馆。此外，它还是历史上影响最为深远的数学家之一——欧几里得的出生地。

*

我们对亚历山大的了解比对欧几里得的了解多得多——虽然欧几里得对人类文明的长期影响更大，这是可以论证的。如果说数学界有一个家喻户晓的名字，那肯定就是“欧几里得”。尽管我们对欧几里得的生平所知寥寥，但对他的著作却相当熟悉。很多个世纪以来，数学和欧几里得在整个西方世界是一组恰当的同义词。

欧几里得为什么如此著名？有些数学家比他更伟大、更重要，但是在近两千年中，整个西欧学习数学的小学生都知道欧几里得，而且在阿拉伯世界这个较小一点的范围内也是如此。他是有史以来最著名的数学原著之一《几何原本》（*Elements of Geometry*，经常被缩写成 *Elements*）的作者。印刷术发明以后，这本书成了第一批以印刷形式出现的书籍之一，出版了一千多个不同的版本，仅次于圣经。

我们对欧几里得生平的了解只比对荷马的了解多那么一点。他公元前 325 年前后生于亚历山大，死于公元前 265 年前后。

刚说出了上面的话，我就不安地意识到自己应该把这些话收回。有人认为确实有过欧几里得这个人，而且是《几何原本》的惟一作者，其实，这只是三种不同推测中的一种而已。第二种推测是，确实有过欧几里得这个人，但是他并没有写过《几何原本》这本书，至少他不是独自写成了这本书。他可能只是一个数学家小组的领导者，是这个小组共同完成了《几何原本》。第三种推测的争议很大，但是有着很大的可能性，那就是确实有过这样一个小组，但是这个小组与 20 世纪中叶那群被统称为“尼古拉·布尔斯基”（Nicolas Bourbaki）的以法国人，多数是年轻人为主的数学家小组很相似，“尼古拉·布尔斯基”只是他们的总笔名而已。但是这其中最有可能的是，欧几里得这

个人确实存在，而且独力创作了《几何原本》。

但这并不是说这本书中的所有内容都是欧几里得一个人的发现。他只是收集整理了古希腊的数学知识。他窃取了前辈们的知识，为后人留下了一笔丰厚的遗产，并借以确立了自己在这一学科中的权威。《几何原本》一般被视为一本几何学著作，但是也有关于数论和某种原型代数（Prototypical algebra）的内容，但这些都被披上了几何学的外衣。

欧几里得的生平我们知道得很少。后来的诠释者们的著作中有一些关于他的零星信息，但是都没有得到现代学者的证实。他们认为欧几里得在亚历山大教书，并据此推测他就出生于该城，而我们对此却并不清楚。在他去世 700 年后，公元 450 年的一本对欧几里得数学的详尽的评注中写道：

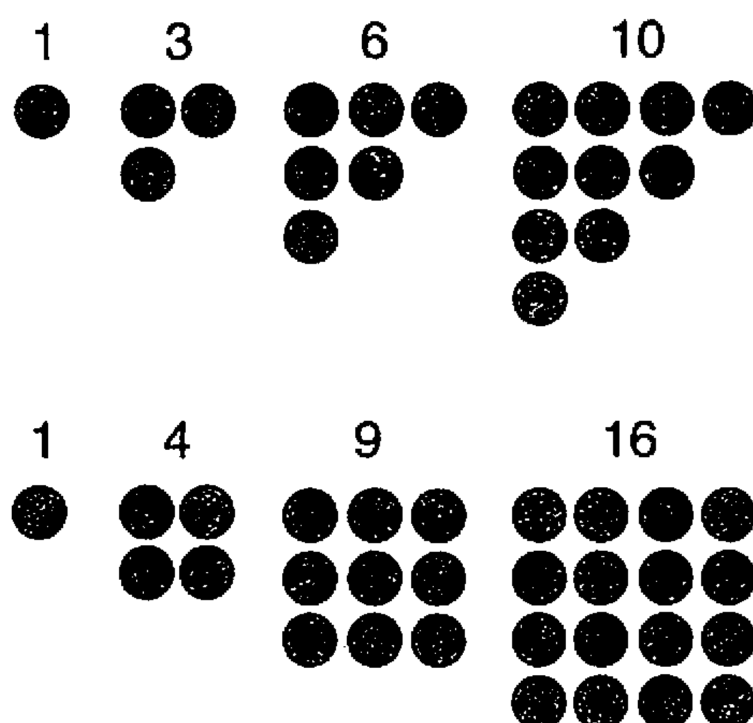
> 欧几里得……汇编了《几何原本》，整理了很多欧多克修斯（Eudoxus）的定理，完善了不少泰阿泰德（Theaetetus）的定理，并且对那些只经过前辈们松散证明的东西进行了不可反驳的论证。欧几里得生活在托勒密一世时期，托勒密一世的贴身随从阿基米德等人曾经提到过他。他们还提到，托勒密一世曾经问他有没有比《几何原本》更好的学习几何学的捷径，而他回答说几何学中没有皇家大道。因此他应该比柏拉图等人年轻，但是比埃拉托色尼（Eratosthenes）以及阿基米德等人要年长。他们都是同时代人，埃拉托色尼曾经在一些地方提起过。欧几里得骨子里是一个柏拉图主义者，赞同他的哲学思想，这就是为什么他把《几何原本》全书的末尾部分以所谓的柏拉图图形（Platonic figure）呈现。

《几何原本》对某些问题的处理方法间接然而明显地证明，欧几里得在某种意义上说应该是雅典柏拉图学园的学生。举例来说，只有在那里，才能学到欧多克修斯和泰阿泰德的几何学。对于他的性格，我们只能从帕普斯（Pappus）的著作片断中了解，他这样评价欧几里得："非常公正，并且很喜欢那些不论通过何种方法提高数学水平的人，为人谨慎，少有愠色，虽说是一个当之无愧的学者，但从不骄矜自恃。"斯托比亚斯（Stobaeus）讲述过一个流传下来的轶事。欧几里得的一个学生问他学习几何能够得到什么，他就叫来他的奴隶说："给他一枚硬币，因为这个人想通过学习获利。"

*

希腊人对待数学的态度与巴比伦人和埃及人有着很大的区别。后两者大体上都将数学看成一种实用性的东西，尽管有时候这种"实用"只不过意味着校准金字塔的轴线，以使死去的法老的灵魂能够得到天狼星的指引。但对于希腊的数学家来说，数字不只是偶尔用来支持他们的神秘信仰，而是数学就是他们信仰的最核心。

亚里士多德与柏拉图谈到了以毕达哥拉斯为中心的教派，这个教派活跃在公元前550年前后，他们把数学，尤其是数，看做万物的基础。基于弦乐器上音符的和谐，他们提出了宇宙和谐的神秘思想。如果一根弦产生一个确定的音，那么一根一半长度的弦就会产生一个高八度的音符，这是最和谐的音程。他们研究了大量的数字组合样式，尤其是用物体排列成的多边形样式。比如说"三角数"（trianglar numbers）1，3，6，10是由三角形组成的；"正方数"（square numbers）1，4，9，16是由正方形组成的，如图：



三角数与正方数

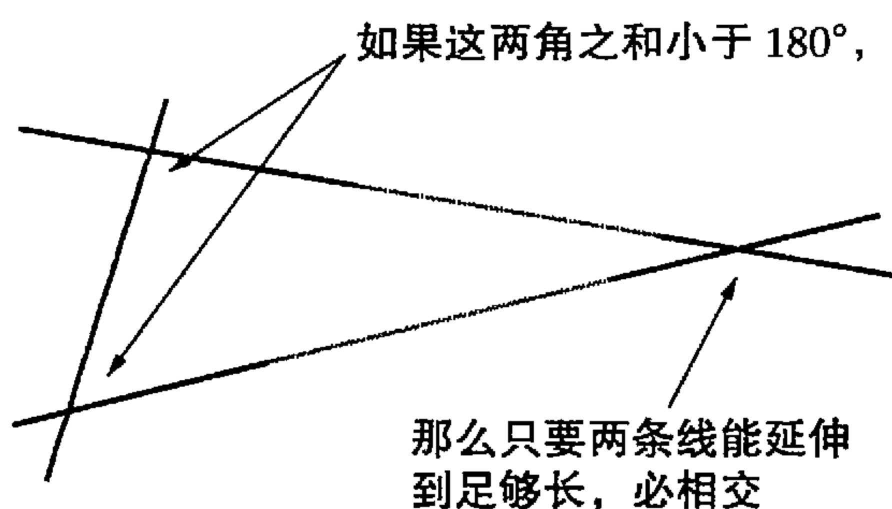
毕达哥拉斯主义中包含某种狂热的数字命理学——如认为 2 是雄性的，而 3 是雌性的——但是这种认为自然的深层结构具有数学性的观点一直延续至今，被当做大多数理论科学的基础。虽然后来的希腊几何学不再那么神秘，但是希腊人通常还是认为数学的终极目标就是它自身，属于哲学的一个分支而不仅仅是一件工具。

我们有理由相信，那时的情况并非完全如上所述。我们非常确定，阿基米德是欧几里得的学生，他用他的数学才能设计了强有力的机械和战争器械。极少数遗留下来的复杂的希腊机械设计巧妙、制作精细，揭示了一种高度发达的技艺传统——“应用数学”的古代形式。最有名的例子也许要数在安提凯特拉岛（Antikythera）附近发现的天文测算设备了，它是用齿轮错综复杂地组装起来的。

欧几里得的《几何原本》很符合古希腊数学的这种纯净观念——也可能是这种纯净观念缘自《几何原本》。这本书的重点是逻辑与证明，而没有实际应用的痕迹。对我们这本书来讲，最重要的不是《几何原本》里写到的东西，而是它所没有写到的。

*

欧几里得作出了两项伟大创新。第一个就是证明的观念。所有不是从已被证明为真的命题开始，经过一系列逻辑步骤推论出来的数学陈述，他全都不接受。第二个创新是，他发现，所有的命题都是从一个初始命题开始的，然而这些初始命题却无法证明。所以他列出了五大假设，作为推论的基础。其中四个是简单易懂的：两点之间可成一线，任何有限长度的直线都可被延长，任意中心和任意半径皆可作一圆，凡直角都相等。



欧几里得的第五条假设

但是第五条假设却非常特殊，冗长而复杂，其结论既不太合理又不太明确。最为复杂的是平行线现象——永远沿着同一方向延伸，然而永不相交的直线，中间距离总是相等，就像一条无限延长的笔直的大路两边的人行道。欧几里得最后认为，两条线只要与第三条线相交，那么，前两者一定会在它们与第三条线形成的两个内夹角之和小于两直角之和的一端相交。这个假设从逻辑上相当于，一条给定的直线必有一条通过某一点的（此点不在给定的直线上）平行线。

许多世纪以来，第五条假设一直被看成一个瑕疵。人们认为，第

五条假设要么可以通过从前面四个假设开始的推论将其驳倒，要么可以代之以一个像其他假设一样明确而又较简单的东西。到了 19 世纪，数学家们才明白欧几里得的第五条假设是完全正确的，因为他们可以证明这个假设根本无法从其他几个假设推论出来。

*

对欧几里得来说，逻辑证明是几何学的本质特征，也一直是数学的基础。对一个缺乏证明的命题是应该存疑的，不论有多少相关证据的支持，也不论它有多么重大的意义。物理学家、工程师和天文学家总是鄙夷逻辑证明，他们觉得这是一种书生气的东西，因为他们有一个更实效的替代品：观察。

比如，如果一个天文学家在测算月球的运行时草草写出了月球运动的数学方程，就会很快陷入窘境，因为看起来并没有精确的解方程的办法。因此，天文学家会对方程添添改改，并且引入大量简化的近似值。数学家却担心这些近似值会严重影响最后的结果，并总想确证它不会出问题。天文学家有一套完全不同的方法，来检验自己结论的合理性，他可以看一下月球的运动是否符合自己的推算结论。如果符合，那就等于说证明了这种方法是正确的（因为得出的结果是正确的），同时也验证了这一理论（因为同样的原因）。这种逻辑是非常干脆的，因为如果一种方法存在数学问题，那就几乎可以肯定它无法预测月球的运动。

没有观察和设备这些奢侈品，数学家只能通过内在逻辑验证他们的理论。一个论点的意义越重大，就越需要逻辑证明。所以，当人们都希望一个论点是正确的，或者如果它是正确的就会产生重大意义的

时候，逻辑证明就更加重要了。

证明不能空穴来风，也不能无限地向后回溯到先前的逻辑。它必须开始于某一点，这一点必须是某些未被证明——也不可能被证明——的东西。我们今天把这种未被证明的起始点叫做公理，这些公理对某种意义上的数学问题来说就是其游戏规则。

如果谁反对公理，就尽可以改变它，但那就是另外一种游戏了。数学决不会断言某一论点为真，它只断言，如果我们作出大量假设，那么与其相关的陈述就必须是一种逻辑推论。这并不是说人们不能挑战公理。数学家们会为了某种目的而争论某一既有公理体系是否优于另一公理体系，也会争论这一体系是不是具有某种本质性的优点或好处。但这种讨论与任何特定公理游戏的内在逻辑无关，而关乎到底哪种游戏更有价值，更有趣，更好玩。

*

欧几里得的公理推论——他长长的、披沙拣金式的逻辑推导链——有着至关重要的意义。比如，他以他那个时代毫无瑕疵的逻辑证明，如果你认同他的公理，就必然可以推出如下结论：

- 在一个直角三角形斜边上的正方形的面积与在此三角形的另外两条边上的正方形的面积之和相等。
- 质数（prime numrbers）有无限多个。
- 存在无理数——它不能用分数准确地表示。比如 2 的平方根。
- 精确地说只有五种规则的立体：四面体、立方体、八面体、十二面体和二十面体。

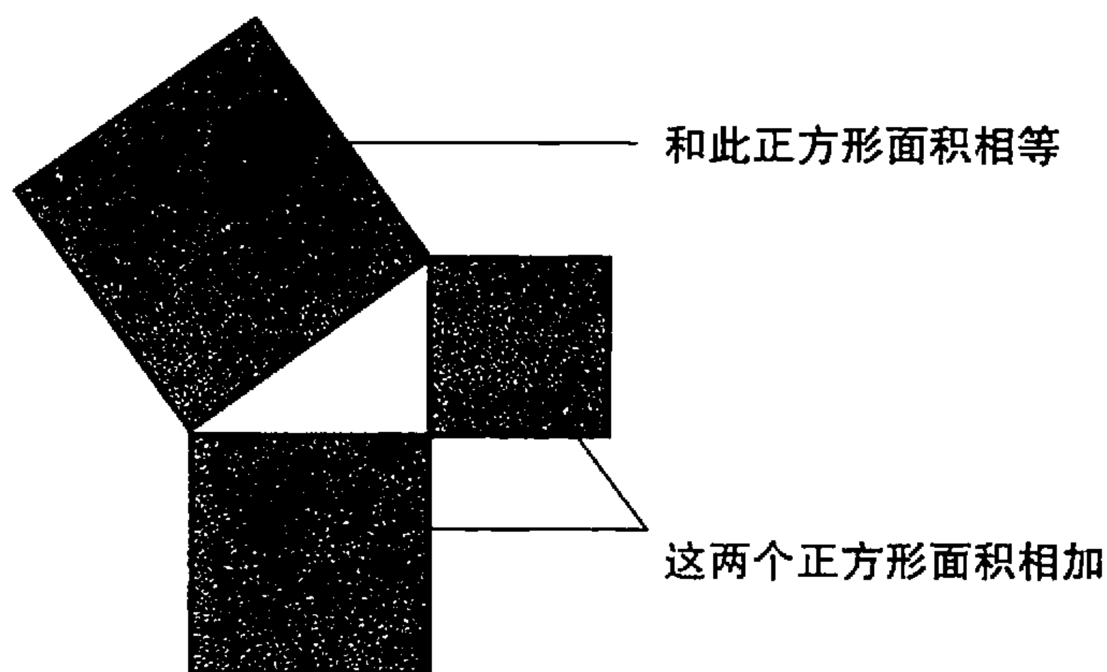
- 任一个角只用直尺和圆规都可以进行二等分。
- 有 3 条、4 条、5 条、6 条、8 条、10 条和 12 条边的规则多边形都可以用直尺和圆规绘出。

我们把得到了证明的数学陈述叫做“定理”。但是欧几里得却与我们的观点大相径庭，他并不直接运用数字。我们解释成数字性质的东西在欧几里得那里都是用长度、面积和体积表示的。

*

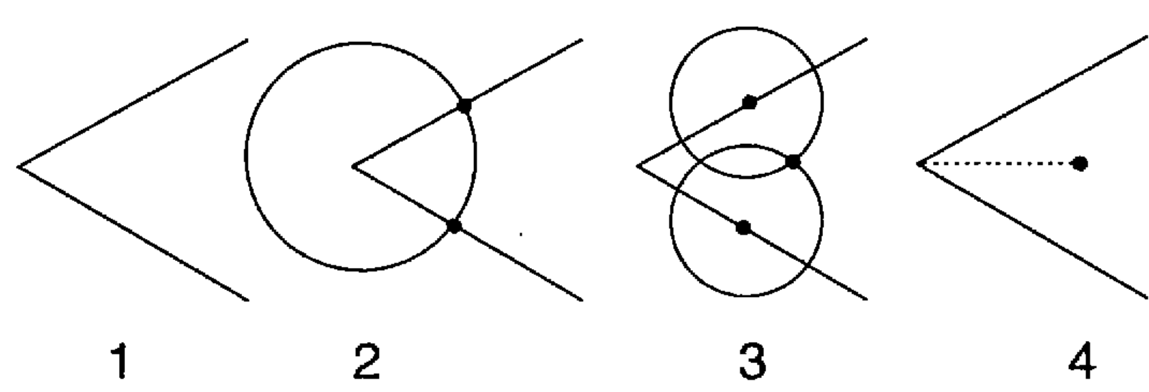
《几何原本》的内容主要分为两大块。一部分是定理，告诉我们某些东西是正确的；另一部分是作图，告诉我们如何去做某些东西。

一个典型而著名的定理是《几何原本》第一卷的命题 47，一般被称作毕达哥拉斯定理。这条定理告诉我们直角三角形的最长边与另外两条边有一种特殊的关系。但是，如果不经过引申和解释，它就一点用处也没有。



毕达哥拉斯定理

第一卷的命题 9 对本书来说很重要，在这一命题中，欧几里得解决了角的“等分问题”。欧几里得等分角的方法简单而巧妙，尽管其初始阶段受到了技术条件的限制。



如何用直尺和圆规等分一个角

用两条线段构成一个角，以两线段的交点为圆心用圆规画一个圆，这个圆就与这两条线段分别相交于两点（黑色圆点）。再分别以这两点为圆心，画两个半径相同的圆。此两圆相交于另外两点（此处只标出一个），我们所求的平分线就同时穿过这两个点。

照此步骤重复作图，就可以把一个角四等分、八等分、十六等分——每做一次，数字就增加一倍，这样，我们就得到了 2 的幂，即 2，4，8，16，32，64，如此等等。

*

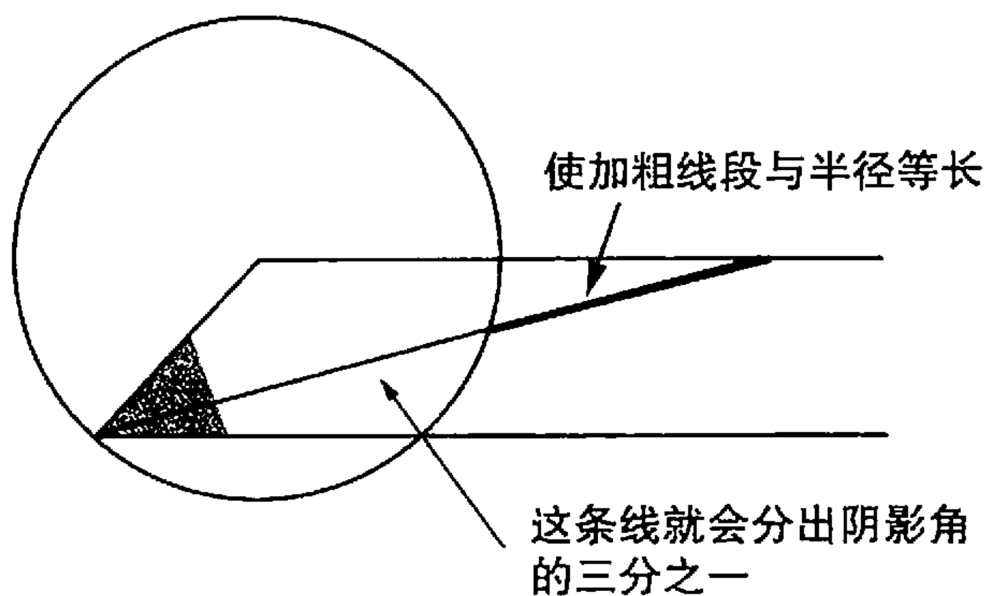
我已经说过，《几何原本》与本书相关的，不是它所写到的，而是它所没有写到的。欧几里得没有告诉我们以下问题的解决办法：

- 把一个角精确地等分成三份（“三等分一角”）。
- 作一个正七边形。
- 作一条长度与一个圆的周长相等的线（“伸直一圆”）。

- 作一个正方形与一个圆面积相等（“化圆为方”）。
- 作一个立方体，使其体积正好是另一立方体的二倍（“倍立方”）。

有人说古希腊人自己就将欧几里得的不朽著作中的这些疏漏之处视为瑕疵，并下了很大工夫对其进行修补。数学史家并没有发现多少支持这一说法的证据。事实上，上述问题古希腊人都能解决，但是他们所用的方法并不在欧几里得这本名著中。欧几里得作图时使用的都是无刻度的直尺和圆规。古希腊几何学家却能使用一种被称为“圆锥截面”（conic sections）的特殊曲线将角三等分；他们还能用另一种被称为“割圆曲线”（quadratrix）的特殊曲线化圆为方。但是另一方面，他们没有意识到，如果可以将角三等分，就可以作一个正七边形（我要强调是作七边形。作一个九边形很容易，但是作一个七边形却是非常巧妙的）。但是，他们明显没有在三等分角之后继续进行深入探究，仿佛他们无心于此。

后来，数家们就开始用一种完全不同的眼光看待欧几里得的疏漏了。他们并没有去寻求解决这一问题的新工具，而是想知道能不能用欧几里得的有限工具——直尺和圆规——解决这些问题（并且不用标有刻度的尺子，使用带刻度的尺子的作图法被希腊人称为纽西斯作图法 [neusis construction]，这种方法可以精确有效地将角三等分，其中有一种方法是阿基米德发明的）。发现什么可以做到，什么无法做到，并给出证明，这需要很长的时间。直到 19 世纪晚期，我们才终于明白，仅用直尺和圆规的话，上述问题一个也解决不了。



阿基米德如何三等分一个角

这是一个重要的进步。数学家们没有去证明某一特定方法可以解决某一特定的问题，恰恰相反，他们试着去证明某某方法不能解决某某问题。他们开始理解数学中固有的局限性。他们正是从让他们发疯的困扰中界定了这种局限性，并证明它确实是不折不扣的局限性。

*

为了避免让大家误入歧途，我将指出三等分角这一问题的若干重要性。

在理想化了的希腊几何公式严苟的条件下，精确作图是个严峻的挑战，在希腊人那里，线要无限细，点是零尺寸。要将一个角精确地三等分，不是精确到十分之一、百分之一或十亿分之一，而是必须做到无限精确。但是即使在这种精神的要求下，我们仍可以把圆规的圆心固定在一个给定的或即将作出的点上；我们可以在两个这样的点之间无限精确地确定圆规的半径；还可以画一条精确地穿过这样两个点的直线。

在杂乱不堪的现实中不会发生这样的事。难道因此就可以说欧几

里得几何在现实世界中毫无用处吗？事实并非如此。比如，如果你照欧几里得的命题9作图，用现实的圆规和现实的纸张，就能作一条漂亮的平分线。在计算机绘图发明以前，人们就是这样绘图的。理想化并不是一种缺陷，而是数学的根本前提。只有在理想化的模式下，才能进行逻辑推论，因为这样我们才能准确地得知对象的性质，但在凌乱的现实世界中却做不到。

但是理想化也有自己的局限性，而且有些时候还会很不适用。在筑路工画一条巷道时，无限细的线就不能用。但这种模式自有其用武之地，它可以帮助我们从几何陈述中得出逻辑结果。作为回报，它还可以帮助我们理解现实世界，然而这绝不是欧几里得思想的中心。

下面这种观点与上述问题相关，但却指向一种截然不同的方向。你在百分之一或十万分之一的误差下，将一个角近似地三等分并不难。由铅笔线的粗细造成的千分之一的误差在技术性作图中是不要紧的。但数学的问题却是进行理想的三等分。任意一角都能够精确地三等分吗？答案是否定的。

人们常说“你可以给出反证啊”，数学家们却认为这是在浪费口舌。另外，反证有着独特的魅力，特别是在人们寻找新的证明方法的时候。反证法通常比正面的解决办法更有效，也更有趣。如果一个人发明了一种新方法，能区别哪些能用尺规作图法来做，哪些不能用尺规作图法，那你就获得了一条全新的思路，还有新的数学理论和工具。

一个人无法使用未创造出来的工具。你不能在电话发明之前给朋友打电话；如果农业没有发明或者火未被发现，你也吃不上苹果派。所以，制造工具和解决问题一样重要。

*

等分角与另一种更加有趣的事情——作正多边形——紧密相关。

多边形就是直线组成的封闭图形（古希腊人称之为“多角形”）。三角形、正方形、长方形、◊状的菱形，都是多边形。圆不是多边形，因为它的边是曲线而不是直线。如果一个多边形的各边长度相等并且每组邻边相交组成的角的度数相等，那它就是正多边形。以下分别是有着 3，4，5，6，7，8 条边的正多边形：



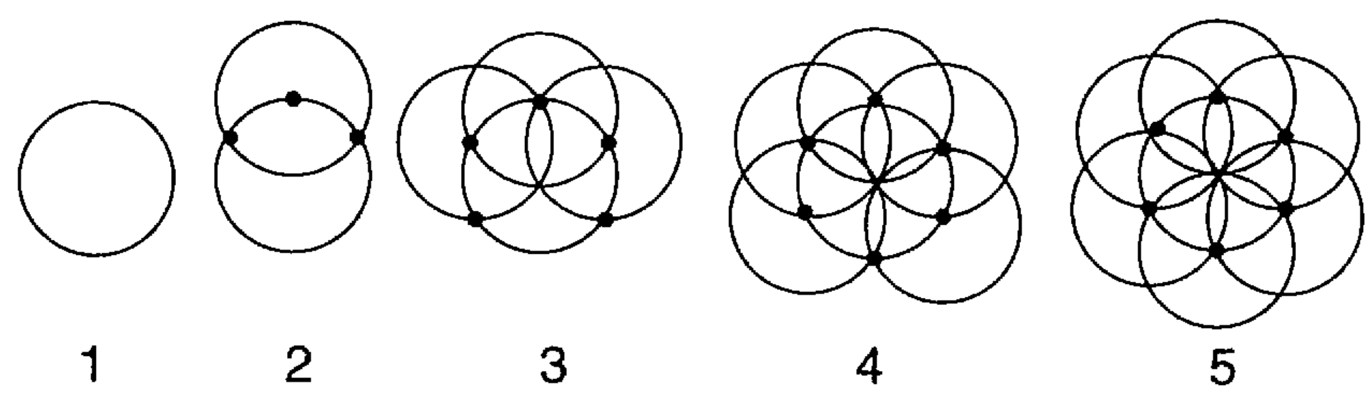
正多边形

它们的学名分别叫等边三角形（equilateral triangle）、正方形（square）、“正”五边形（[regular] pentagon）、六边形（hexagon）、七边形（heptagon）和八边形（octagon）。这些名字有时候也被不太讲究地写成 3-gon，4-gon，5-gon，6-gon，7-gon 和 8-gon。这些术语看起来很不够雅观，但是，在表示正十七边形（我们马上就会讲到）时，“17-gon”就比“heptadecagon”或“heptakaidecagon”实用得多了。当你碰到 65537 边形（65537-gon）时（确实有这种情况），那就更有体会了。

关于哪些正多边形可以作出来的问题，欧几里得和他的先辈们肯定思考了很多，因为他们提出了很多此类图形的作图方法。这是一个令人着迷又着实棘手的问题。古希腊人知道如何构造边数为 3，4，5，6，8，10，12，15，16，20 的正多边形；我们现在知道边数为 7，9，

11，13，14，18，19 的正多边形是无法被作出的。这些数字中，惟独没有出现 17。我们会在合适的时候讲到正十七边形，它的重要性是多方面的，而不仅仅是在数学方面。

在几何作图的过程中，仅有的工具就是现实的纸张、现实的直尺和现实的圆规。这让你好奇，这些玩意儿要如何整合起来呢？那我就来告诉你我所钟爱的正五边形的作图方法吧。我是从我叔叔在 20 世纪 50 年代末期给我的一本书中学会这种作图方法的，这本让人爱不释手的书叫《人必须懂得度量》（*Man Must Measure*）。



怎样作出一个正五边形

从头到尾都使用同一半径，这样所有的圆的大小就会相等。

（1）作一圆；（2）在此圆的圆周上选择一个点，以之为圆心再作一圆，这个圆就会与初始圆相交于两点；（3）再以这两点为圆心作圆，得到两个新的交点；（4）以这两个新的交点为圆心再作两个圆，它们会同时穿过一个与初始圆形成的新交点。连接这六个点就会得到一个正五边形。如果完成（5）：以这第六个点为圆心再画一个与初始圆两点相交的圆使之形成一个花朵的形状，就会颇具美感，但这在数学上并不必要。

欧几里得的方法与这个方法很相似，更简单，但是没这么漂亮，他还证明了其方法的可行性。你可以在《几何原本》的第四卷的命题 15 中找到这种方法。

波斯诗人

醒来吧！太阳驱散了群星

暗夜从空中逃遁，

灿烂的金箭

射中了苏丹的高瓴。*

对我们大多数人来说，“奥玛尔·海雅姆”（Omar Khayyám）这个名字一直是和讽刺长诗《鲁拜集》（*Rubaiyat*）联系在一起的，尤其是爱德华·菲茨杰拉德（Edward Fitzgerald）优雅的英译本。但是对数学史家来说，海雅姆声名显赫还有一个更重要的缘由。在西欧的学者们坠入黑暗时代，在宗教的喝斥面前放弃了他们的定理—证明思想的时候，擎起古希腊数学家丢掉的火把，继续推动数学向前发展的波斯和阿拉伯数学家当中，海雅姆是卓然超群的一位。

---

* 本书所引用的《鲁拜集》中的诗均出自郭沫若的译本，只是按照现在的语言习惯进行了少许修改。如其译本中本诗第一行起始为“醒吧”改为“醒来吧”。——译者注，下同

在海雅姆的伟大成就中，最了不起的是他通过古希腊几何学的卓越技巧解决了三次方程问题。他的技巧十分必然地超越了默默地限制着欧几里得几何学的尺规作图。原因就是这些工具是不称职的，希腊人对这一事实颇有察觉，但却无法证明，因为他们缺乏一个必要的视角。这一视角不是几何学，而是代数学。但是海雅姆的方法并没有超越尺规作图太远，他还要依赖于那种被称为“圆锥曲线”的特殊曲线，因为他们可以通过用平面切割圆锥体作图。

*

科普读物的传统智慧是：每一个方程都会让一本书的销量减少一半。如果真是这样，可真不是什么好事。因为如果没有方程来演示，那就没人能够理解这本书的关键论题了。比如下一章，是讲文艺复兴时期的数学家在解三次方程和四次方程上的发现的。我可以不向你描述解四次方程的公式是什么样的，但是，我们真的需要至少粗略地浏览一下解三次方程的公式。否则，我能告诉你的就只有如下等等的玩意儿了：“用某些数字乘以另一些数字，再加上几个数字，然后就开平方，再加上一个数字然后开立方；然后换一些不同的数字重复同样的过程，最后把这两个结果相加。天哪，我差点忘了——你还得做点儿除法。”

有些作者挑战了这种传统智慧，他们甚至就写关于方程的书。他们都遵从了一句古老的俏皮话：“如果你装了一个假肢，就用它打打招呼。”现在，这本书感觉就是关于方程的。但是就像你要写一本关于山的书不该要求你的读者都去爬山一样，写一本关于方程的书也不能要求读者非得解方程不可。但我还是要说，如果关于山的书的读者从来没有见过山，他们就无法理解这本书。所以，列举一些我精选出

来的方程，对你我来说都将有很大好处。

我的做法的基本原则是“展示”，而我这么做主要是出于为你的考虑。我需要你去“看”，但不需要你去“做”方程。需要时我会把方程分成几个部分，然后告诉你它的哪些特点与我们的论题有关。我绝不会要你去解一个方程，也不会要你去计算，我会尽力让你避开这些东西。

一旦你理解了方程，就会发现它们非常可亲。它们明晰、简练，有时还很美。我们应该知道，其实方程是表达某种计算“处方”的简单、明晰的语言。我会尽量用词句来表述这些处方，或者让你感觉到它到底是怎么回事。但在极少情况下，如果我发现用词句表述过于麻烦，我也会运用一些符号。

这本书中有三种重要符号，我现在讲一下其中的两种。其一是我们的老朋友 x，就是“未知数”。它代表一个我们尚且不知的数字，但我们正在发狂地想求出它的值。

第二种符号是一些上标的小数字，比如 ^{2}，^{3}，^{4}。我们用它们表示一些数字自身相乘的次数。因此 5^3 就表示 $5\times5\times5$，也就是 125，而 x^2 就表示 $x\times x$，这里的 x 就表示我们的未知数。这些小数字读作“平方”、“立方”、“四次方”，等等，并且被总称为相关数字的“乘方”。

别问它们为什么叫这样的名字，因为它们总得有个称谓。

*

巴比伦人解二次方程的方法对古希腊人来说可能已经过时了，也可能是被改造了。生活于公元前 100 年和公元 100 年之间亚历山大某处的，名叫海伦（Heron）的人，用希腊术语讨论了一个典型的巴比伦风格的问题。公元 100 年前后，尼克马库斯（Nichomachus）——

他可能是一个来自阿拉伯的“犹地阿”（Judea）人——写了一本《算术入门》（*Introductio Arithmatica*），这本书抛开了古希腊用几何量的方式表示数的传统，比如表示长度、面积等。对尼克马库斯来说，数就是独立地代表它自己的量（quantity），而不是线的长度。尼克马库斯是一位毕达哥拉斯主义者，在他的著作中，他只用整数及其比例，而且从不用符号。在接下来的一千年中，他的著作成了算术的范本。

公元500年前后，从丢番图（Diophantus）的著作开始，符号思想进入了代数学。关于丢番图，我们知道的惟一一件事情就是他死时的岁数，而且是通过一个不大可信的途径得知的。一本希腊的代数问题汇编中这样写道：“童年占了丢番图六分之一的生命；随后的十二分之一是他的青春期；又过了七分之一，他结了婚，五年后他的儿子出生；他这个儿子活了父亲寿命的一半，父亲比儿子晚死了四年。丢番图死的时候是多大年龄？”

运用古代代数学家的方法，或者加点现代的方法，就可以算出丢番图死时是84岁。这个岁数可算得上是长寿了，但是这个代数问题所依据的事实却很可疑。

这就是我们知道的关于丢番图生活的一切。但是，通过后来的抄本以及另一些文献中的引述，我们对他的著作了解得很多。他写了一本关于多角数的书，其中有一部分流传了下来。这本书是用欧几里得的体例写成的，他用逻辑推论证明定理，并没有多大的数学意味。《算术入门》中有13卷比较有数学意味。多亏了13世纪一个对更早抄本的希腊文抄本，其中的六卷才得以保存至今。另外有四卷可能也已经在一批发现于伊朗的手稿中浮出水面，但是，并不是所有的学者都相信它是丢番图的作品。

《算术入门》提出了一系列的问题。丢番图在前言中说，他写这

本书是为了给他的一个学生作练习。他用一个特殊符号表示未知数，并且用不同的符号，看上去是dynamis（乘方）和kybos（三次方）这两个单词的缩写，来表示这个未知数的二次方和三次方。丢番图的符号体系并不完善，他通过相邻排列表示符号的相加（就像我们现在表示相乘一样），但用一个特殊符号来表示减法。他甚至有一个表示相等的符号，但也有可能是后来的抄录者添加进去的。

《算术入门》大体上是讲解方程的。流传下来的部分的第一卷讨论了线性方程，其余的五卷涉及了二次方程的很多问题，一般都有好几个未知数，还讲到了某些特殊的三次方程。这些方程的一个关键特征是，其结果都是整数或有理数。我们今天把其结果仅限于整数和有理数的方程称为“丢番图式的”。《算术入门》中有一个典型例题：“找出三个数，使它们的和以及任意两个的和都是一个完全平方。”我们试试看——这绝非易事。丢番图的答案是41，80和320。这三个数的和是$441=21^2$，任意两个相加是：$41+80=121=11^2$；$41+320=361=19^2$；$80+320=400=20^2$。真是妙不可言。

丢番图方程在现代数学中非常重要。一个著名的例子就是费马的“最后定理”，这条定理认为：两个完全的三次方数之和不可能等于另一个三次方数，两个更高次乘方数之和也不可能等于另一个更高次乘方数。对二次方来说，这很简单，毕达哥拉斯就已经解决了这一问题，如：$3^2+4^2=5^2$或$5^2+12^2=13^2$。但是三次方就不行，四次方、五次方或者其他高于二次方的都不行。1650年，皮埃尔·德·费马（Pierre de Fermat）在自己的那本《算术入门》的页边上记下了这一猜想（未经证明，名为定理但名不副实）。将近350年后，生活于美国的英国数论家安德鲁·威尔斯（Adrew Weils）证明了费马这一猜想的正确性。

数学传统的历史接续有时需要很长的时间。

*

代数真正获得数学意义可以追溯到公元830年，这时数学活动的中心已经从希腊世界转移到了阿拉伯世界。那时候，天文学家穆罕默德·阿尔·花拉子模（Mohamed ibn Musa al-Khwārizmī）写了一本名为《移项和消去的科学》（*al-Jabrw' al Muqâbala*）的书，可粗略地翻译为“复位和简化”。这组单词表示一些处理方程以便以更好的形式求解的标准技巧。我们的“代数”（algbra）一词就是由“al-jabr”得来的。它在12世纪的第一个拉丁译本的标题是“*Ludus Algbrae et Almucgrabalaeque*”。

花剌子模的著作中有受之前巴比伦和希腊影响的痕迹，并且也有赖于公元600年前后印度的婆罗摩笈多（Brahmagupta）所提出一些观念。书中解释了线性方程和二次方程的解法。花剌子模的直系弟子发现了解决一些特殊的三次方程的解法。其中一位是伊本·瓜拉（Thâbit ibn Qorra），他是医生、天文学家和哲学家，是生活在巴格达的一名异教徒；还有一位名叫阿勒-哈森·伊本·阿勒-海赛姆（Al-Hasan ibn al-Haitham）的埃及人，在之后的西方著作中通常被叫做Alhasen。但是他们当中，最著名的还是奥玛尔·海雅姆。

奥玛尔的全名是吉亚斯·阿尔丁·阿布·法斯·奥玛尔·本·伊卜拉欣·内沙布里阿尔-海雅姆（Ghiyath al-Din Abu' l-Fath Umar ibn Ibrahim Al-Nisaburi al-Khayyámi）。“al-Khayyámi”的字面意思是“制作帐篷的人”，有些学者认为这可能是他父亲伊卜拉欣的职业。奥玛尔1047年生于波斯，在纳霞堡（Naishapur）度过了他高产的一生。这个地方属于伊朗北部与土库曼斯坦接壤的呼罗珊（Khorosan）省，

离马什哈德（Masshad）很近。

传说奥玛尔年轻时离开家跟从著名的纳霞堡牧师伊曼·穆瓦法克（Iman Mowaffak，波斯大学问家）学习伊斯兰教义和可兰经。他与两个同窗，哈桑·萨巴赫（Hasan Sabah）和尼让牟·穆勒克（Nizam al-Mulk）缔结了友谊。他们三个人立下誓约，如果他们当中任何一个人富贵了——作为穆瓦法克的学生，这很有可能——就要与另外两个人同享他的财富和权力。

几年后，这几个学生完成了学业，但誓约继续有效。尼让牟去了喀布尔，海雅姆没有那么大的政治野心，做了一段时间的“帐篷制作人”，这也可能是他的名字“al-Khayyámi”的另一来源。科学和数学成了他的至爱，并把大部分的业余时间都花在了这两样东西上面。最后，尼让牟回来了，在政府中谋得了一个职位，为阿尔普·艾尔斯兰（Alp Arslan）苏丹管理政事，在纳霞堡拥有一座官署。

由于尼让牟富贵了，哈桑和奥玛尔就去兑现誓约。尼让牟请求苏丹援助他的两个朋友，获得同意之后，他履行了誓约。哈桑获得了一个俸禄优厚的公职，然而奥玛尔却只请求继续他在纳霞堡的科学研究，他会在那里为尼让牟祈祷健康和安宁。他的老朋友为奥玛尔安排了政府薪俸，让他腾出时间进行研究。

后来，哈桑因试图扳倒上级官员而丢掉了自己的闲职，平心静气的奥玛尔却按部就班地晋升到了改革历法的班子中。波斯历法以太阳的运行为基础，新年的第一天总是变来变去，非常混乱。这个工作需要一个称职的数学家，需要奥玛尔运用他的数学和天文学的推算知识，算出指定年份的新年的第一天。

这段时间前后，他还写了《鲁拜集》，不严格地翻译出来就是“四行诗”。这是一种诗体，每首诗都是通过独特的韵律组织起来的四

行韵文，准确地说，是每两行押一韵。《鲁拜集》就是这种诗的集本，其中一首诗就是对其历法改革一个明晰的诠释：

啊，人们说我的推算高明，
我曾经把旧历的岁时改正——
谁知道那只是从历书之中
消去未生的明日和已死的昨晨。

奥玛尔的诗有着鲜明的非宗教色彩。他的很多诗都颂扬酒和它的功效：

在日前，茅店之门未闭，
黄昏中闪现天使的形体
肩着一个土瓶，她叫我尝尝
土瓶里原来是——葡萄的酒浆！

也有对酒的挖苦嘲讽：

莫问是在纳霞堡或在巴比伦，
莫问杯中的是苦汁或是芳醇，
生命的酒浆滴滴浸漏不已，
生命的绿叶叶叶飘堕不停。

他还有些诗揭露了宗教信仰的闹剧。有一首诗质疑了苏丹让某人当自己的随从的想法，以及伊玛目（imam，穆斯林政教领袖）对自己

讲道成果的想法。

其间，颜面尽失的哈桑被驱逐出了纳霞堡，加入了一伙强盗，并利用自己所受的高等教育当上了他们的头头。1090年，这伙强盗在哈桑的指挥下夺取了里海南岸厄尔布尔士山（Elburz Mountain）上的阿拉穆特（Alamut）城堡。他经常掠扰这一地区，变成了臭名昭著的“山中老人”。他的喽啰们因为吸食麻药（hashish，一种非常剧烈的大麻产品）被称为“吸麻药的”（Hashishiyun）。他们在山中建了六个要塞，经常从里面蹿出来杀害他们选定的宗教和政治人物。他们的名字就是“刺客”（assassin）一词的来源。就这样，哈桑作为穆瓦法克的学生，用自己的方式获得了财富与名声，虽然那时他已不愿与老同学同享富贵。

当奥玛尔推算出天文表并找到三次方程的解法的时候，尼让牟的政治生涯结束了，他极具讽刺性地被哈桑的徒众暗杀了。奥玛尔活到了76岁，据说，他死于1123年。哈桑死于第二年，时年84岁。刺客们的暗杀不断地制造着政治浩劫，直到蒙古大军在1256年占领了阿拉穆特，才把这伙人荡平。

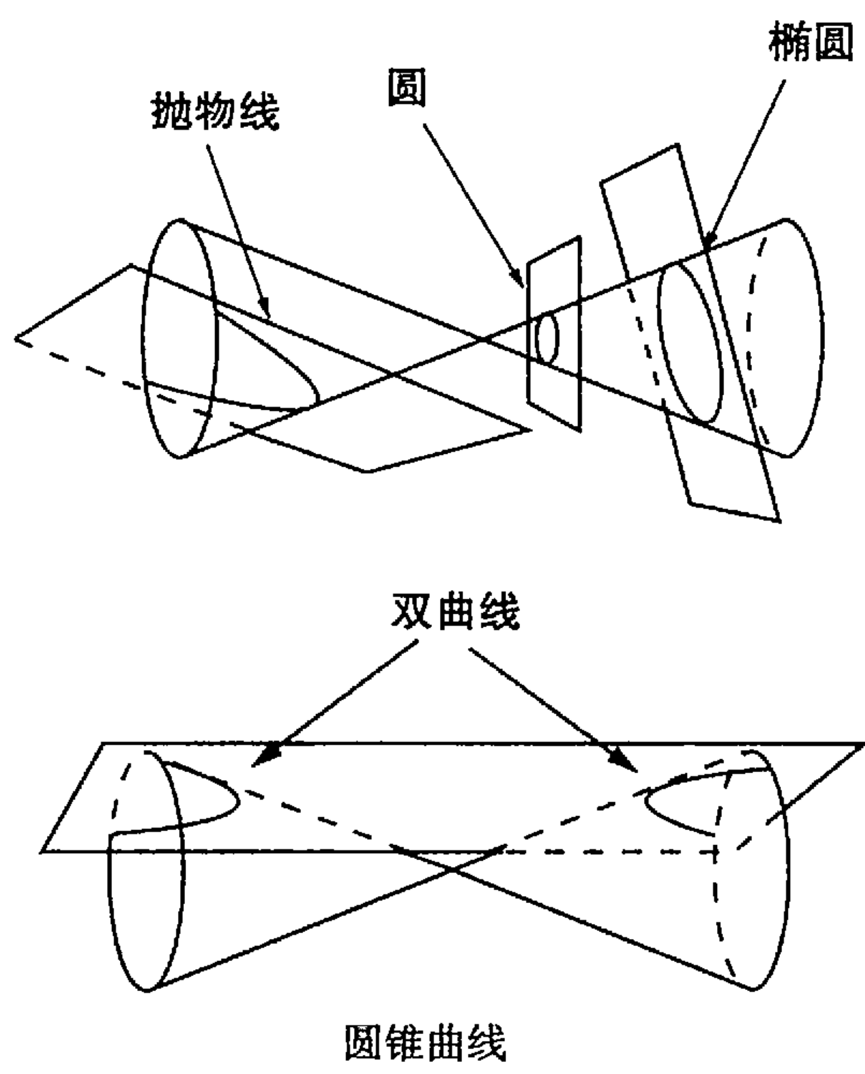
*

让我们回到奥玛尔的数学成就上来。公元前350年前后，希腊数学家门奈赫莫斯（Menaechmus）发现了一种特殊曲线——圆锥曲线。学者们认为，他是要用这种方法来解决倍立方问题的。阿基米德提出了这种曲线的基本原理，佩尔格的阿波罗尼斯（Apollonius of Perga）在《圆锥曲线》（*Conic Sections*）一书中将之系统化，并且扩展了这一论题。让海雅姆感兴趣的是，圆锥曲线可以用来解决某些三次方程。

这种曲线被称为圆锥曲线是因为，它可以通过用一个平面切割圆

锥获得，更恰当地说，是一个双圆锥体，就像两个尖端相连的圆筒冰激凌。一个圆锥体是由相交于同一点并穿过某一圆的线段集合构成的，这个圆就是圆锥的“基部”。但是在希腊几何学中，你可以随心所欲地延长线段，因此就会形成双圆锥。

圆锥曲线有三种主要类型：椭圆、抛物线和双曲线。椭圆是一种卵形封闭曲线，当切面只切割双圆锥中的其中一个时就会得到（圆是一种特殊的椭圆，切面正好垂直于圆锥的中轴时，就会形成一个圆）。双曲线包括两个对称的开放曲线，同时切割双圆锥的两个部分就可以得到，原则上讲它可以无限延伸。抛物线是一种过渡形式，它只有一条开放曲线，在这种情况下，切面必须平行于圆锥表面上的某一条线。



圆锥曲线

如果距离圆锥体的尖端非常远，那么双曲线就会非常接近于两条直线，它们平行于穿过尖端并将双圆锥平分的平面上的直线。这些非常接近于直线的线叫做渐近线。

希腊几何学家对圆锥曲线广泛的几何研究为超越欧几里得定下的

思路营造了相当大的进步空间。这些曲线对现代数学来讲至关重要，但现代数学的关注点却与古希腊人大异其趣。从代数的观点看，这些曲线之简单仅次于直线，对应用科学来说也十分有用。太阳系中的行星的运行轨道就是椭圆曲线，比如开普勒根据第谷·布拉赫（Tycho Brahe）的观测数据推算出了火星轨道。椭圆形轨道是引导牛顿将其关于万有引力的著名“反平方律”公式化的重要观测数据之一。进而，人们认识到，宇宙从某些方面来说展现出明晰的数学模式。这全面开启了天文学中行星现象的可计算化进程。

*

奥玛尔的主要数学功绩就是方程理论。他对两种方程解法进行过思考。第一种是他从丢番图那里继承下来的，他称其为整数方程 “代数式”解法，但更合适的形容词应该是“算术式”；第二种方程的解法被他称为“几何式”，他认为这种解法可以通过几何意义上的特定长度、面积和体积构建起来。

通过对圆锥曲线的自如运用，奥玛尔提出了三次方程的几何解法，并且在 1079 年用他的“代数”完成了对这种方法的解释。由于那时候的人们还没有注意到负数问题，所以方程的各项都只涉及了正数。这导致人们不得不大量地对情况作出区分，现在来说，我们只要在同一个数字前加上一个表示其性质的符号加以区分就可以了。奥玛尔根据各项在方程中的位置把三次方程划分为 14 种。其分类如下：

立方 = 平方 + 边 + 数字

立方 = 平方 + 数字

立方 = 边 + 数字

立方 = 数字

立方 + 平方 = 边 + 数字

立方 + 平方 = 数字

立方 + 边 = 平方 + 数字

立方 + 边 = 数字

立方 + 数字 = 平方 + 边

立方 + 数字 = 平方

立方 + 数字 = 边

立方 + 平方 + 边 = 数字

立方 + 平方 + 数字 = 边

立方 + 边 + 数字 = 平方

上述每一条中的各项数值系数都必须是正的。

你可能觉得很奇怪，上面为什么不包括像是“立方 + 平方 = 边”的情况呢？因为在上述那些情况中，我们能够用未知数分割方程的任何一边，从而将其简化成一个二次方程。

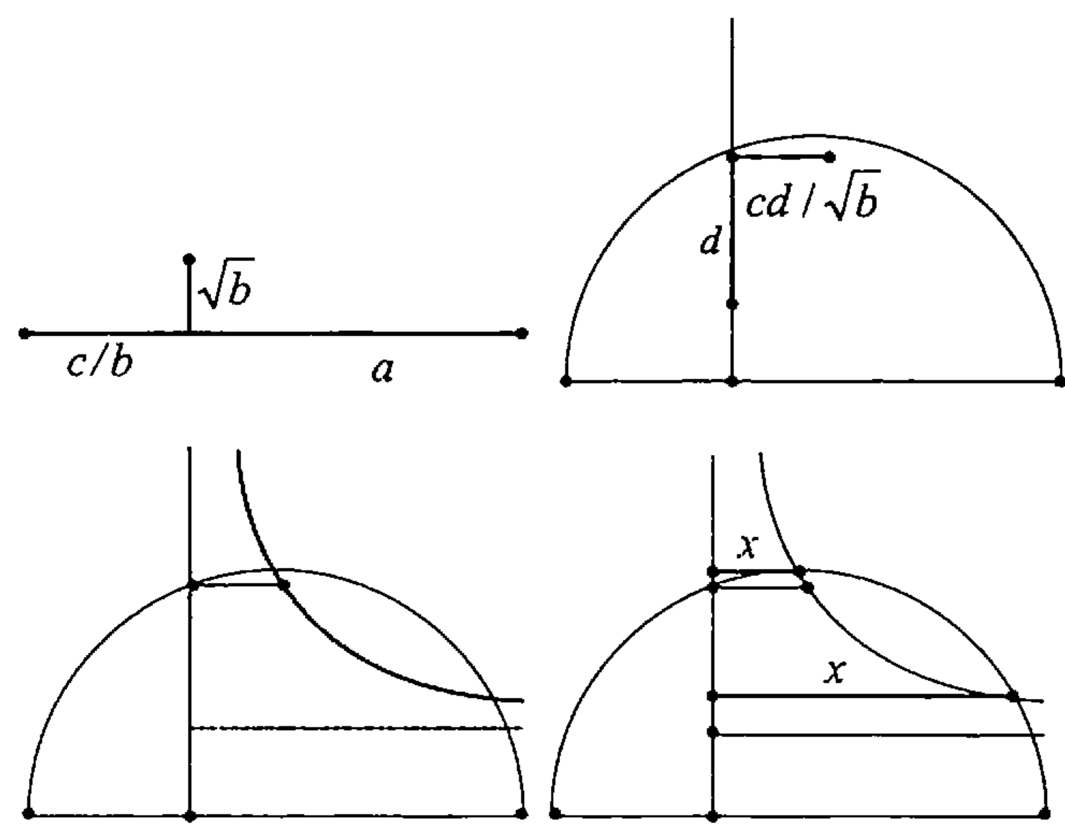
*

这些方程解法并不是奥玛尔的独创，而是建立在先前古希腊人用圆锥曲线解决一些三次方程的方法的基础上的。他系统地发展了这种方法，并用其解决了上述 14 种三次方程。他解释说，以前的数学家已经发现了解很多情形下的三次方程的方法，但是这些方法都很特殊，所解的方程的结构也都限定在各自不同的情况下；还没有人总结

出所有可能情况的方程，更别提解决它们了。“我，相反的——我从未停止期望能精确地知道所有可能的情况，并且从每种情况中区分出哪些可能哪些不可能。”他所说的“不可能”是指“没有正数解”。

将他的成就加以引申，可以说他解决了“一个三次方、一些边和一些数字等于一些平方”的问题。这种问题可以写成 $x^3+bx+c=ax^2$（如果我们不考虑正与负的问题，就能把等号右边的项移到另一边，并把 a 变成 $-a$，于是就可以写成 $x^3-ax^2+bx+c=0$）。

奥玛尔告诉其读者的解方程步骤是这样的：(1) 画三条长度分别为 c/b，$\sqrt{b}$ 和 a 的线，三者形成一个直角；(2) 画一个直径等于水平线段的半圆，将垂直线延长穿过该半圆，如果垂直线的实线部分长度为 d，如图所示画一条长度为 $cd/\sqrt{b}$ 的水平实线；(3) 画一条双曲线（用实线表示）通过刚才形成的点，灰线表示它的渐近线（曲线无限向其趋近的特殊直线）；(4) 找出双曲线和半圆的交点。如图，标注为 x 的两条实线的长度都是三次方程的（正数）解。



奥玛尔的三次方程解法

细节对整体风格来说无关紧要。方法无外乎是用直尺和圆规进行

欧几里得式作图，引入双曲线，然后进行更多的欧几里得式作图——完成。

奥玛尔用类似的作图解决了全部上述 14 种方程，并证明了其正确性。他的分析是有漏洞的：如果系数 a，b，c 的数值不合适，他在作图过程中所需要的点就不存在。比如前述的作图，那条双曲线可能根本无法和圆相交。抛开这种吹毛求疵，他的工作还是卓越而系统的。

奥玛尔诗集中的有些意象是数学性的，而且仿佛暗指了自己的工作，自我否定的语调也贯穿始终：

是与非是虽用几何可以证明，
上与下虽用几何可以论定，
人所欲测的一切之中呀，
除酒以外，我无所更深。

其中有一节尤其引人注目：

我们是活动的幻影之群，
绕着走马灯儿来去，
在一个夜半深更，
点燃在魔术师的手里。

这让人想起了柏拉图关于“洞穴墙壁上的影子”的寓意。这正是对代数和人类境况作出的象征性写照。奥玛尔的天赋使他同时记录下了两者。

嗜赌的学者

"我以上帝神圣的福音书和一个真正的男人的荣誉发誓，如果你把你的发现教给我，我不仅不会把它们发表出去，还要向你保证，作为一个虔诚的基督徒，我会把它们写成密码，这样在我死后也没人能够看懂它们了。"

据说这个神圣的誓言是在1539年发下的。

文艺复兴时期的意大利是创新的温床，数学也不例外。在那个时代革故鼎新的氛围中，文艺复兴的数学家们决心超越古典数学的局限。其中的一位解决了神秘的三次方程。他现在正在控告--位剽窃其秘密者。

这位愤愤不平的数学家就是尼柯洛·冯塔纳（Niccolo Fontana），昵称"塔尔塔里亚"（Tartaglia），是个结巴。那位被称为窃贼的极富才华的家伙是一个数学家、医生、无可救药的无赖和积习难改的赌徒。他的名字是吉罗拉莫·卡尔达诺（Girolamo Cardano），又叫做杰罗姆·卡丹（Jerom Cardan）。1520年左右，吉罗拉莫这个十足的纨绔

子弟，靠着父亲的遗产维持自己的生活。破产之后，他开始以赌博为收入来源。他把自己的数学才能派上了更实际的用场，想借此赢钱。他有一群靠不住的玩伴，有一次，他怀疑一个玩家作弊，就用刀划破了那家伙的脸。

这是一个艰难的时期，但是吉罗拉莫十分顽强。他还是一位很有原创性的思想家，写出了历史上最著名、最有影响的代数学著作之一。

*

我们对吉罗拉莫了解之所以这么多，是因为他在1575年把自己的一切都写进了《我的人生》（*The Book of My Life*）这本书中。书是这样开头的：

> 这本人生之书是按照哲学家安东尼*的先例来写的，我要颂扬人的智慧和美德。因为终有一死，所以也就没有什么完美的人，无论谁都难免为人诟病。因此，任何人生目标，无论多么诱人，都不如认识真理来得有价值。
>
> 我要声明，我写这本书既不是爱慕虚荣，也决不添油加醋。我只写我的切身经历，还有一些为我的学生所知的事，有些事他们还曾参与其中。这是对我一生中的这些横断面的简要记述，是我的人生之书。

和当时很多数学家一样，吉罗拉莫也研习占星学，他解释了自己

---

* 即古罗马皇帝马克·奥勒留。

出生时的星象：

> 我听说母亲用过很多的堕胎药，但都没能把我打掉，我在1500年9月24日正常地出生了，时间就在入夜后半个小时到三刻钟之间。……由于各天体的位置相冲，火星的邪恶影响遍及诸天体，它的方位与月亮正好相对。
>
> ……如果不是水星掌管的处女座与前述星位成29度角，我很可能会是个妖怪。如果处女座和月亮的星运不像当时那样作用于处女座的第二区间上，我就可能已经是一个妖怪了，那时间距离我从娘胎里出来时可真是只差一点啊。
>
> 我就这么出生了，或说被我母亲的暴力方法给弄了出来，差点死掉。我的头发是卷曲的黑发。我沐浴着暖酒而获得重生，要是其他婴儿肯定死定了。母亲整整辛苦了三天才把我生出来。

《我的人生》中有一章罗列了吉罗拉莫的著作，他所列的第一本就是《伟大的艺术》(*The Great Art*)，是他提到的三本关于“数学之契约”的书中的一本。另外，他还写过天文、物理、道德、宝石、水、医药、占卜和神学方面的著作。

在这些书当中，只有《伟大的艺术》与我们的故事有关，它的副标题“代数的法则”也说明了这一点。在这本书中，吉罗拉莫收集的不仅是巴比伦人就已了解的二次方程的解法，还有三次方程和四次方程的新解法。与海雅姆的解法不同，《伟大的艺术》并不仰赖圆锥曲线，而是纯代数的。

*

我之前提到过两种数学符号，比如 x^3 就同时表示出这两种符号。第一种符号用字母来表示数（如这里的 x）——未知数或已知的任意数。第二种是用上标的数字表示幂，因此上标 ^{3} 表示 $x \times x \times x$。现在我们要讲到第三种符号，这是我们需要的最后一种符号。

这第三种符号非常漂亮，它是这样的：$\sqrt{\ }$。这个符号表示“平方根”。比如$\sqrt{9}$，表示“九的平方根”，表示自身相乘等于 9 的数。由于 $3 \times 3=9$，因此$\sqrt{9}=3$。但问题不会总是这么简单。最恶名昭彰的平方根——它是因一个关于梅塔蓬托（Metapontum）* 的希帕索斯（Hippasus）的荒唐传说而引起数学家们的注意的，传说此人从一艘船上被扔了下去——就是二的平方根，$\sqrt{2}$。要想用小数精确地表示这个数字，那就会没完没了地继续下去。它的开头是这样的：1.4142135623730950488，但这并没有结束，因为上面这个数字的平方精确地说是 1.99999999999999999999522356663907438144，并不完全等于 2。

我们现在都知道这个符号的来源。它是字母“r”的变形，代表“radix”，即“根”（root）的拉丁语。数学家们就这样理解$\sqrt{\ }$，并把$\sqrt{2}$读作“根号二”。

立方根、四次方根、五次方根等就在根号前加一个上标的小数字，写做$\sqrt[3]{\ }$，$\sqrt[4]{\ }$，$\sqrt[5]{\ }$。

立方根表示一个其三次方等于某个给定数字的数。因此，8 的立方根就是 2，因为 $2^3=8$。我们也可以用小数近似地表示 2 的立方根，

* 地名，属于大希腊的一部分，现为意大利南部塔兰托附近。据说，毕达哥拉斯就死在那里。

以 1.25992104989487316476 为开头向后延续，如果你有足够的耐心，就可以一直延续下去。

正是这个数字带出了古老的倍立方问题。

*

到了公元 400 年的时候，希腊已经不再是数学的前沿阵地了，而是东移到了阿拉伯、印度和中国。这时，欧洲坠入了“黑暗时代”，虽然没有人们通常描述的那样黑暗，但也黑暗得可以了。基督教的蔓延对教堂和修道院里的集中研习和学术探讨造成了很多负面的影响。很多的僧侣抄录欧几里得等人的数学著作，但几乎没有人知道自己在抄些什么。早期的盎格鲁-撒克逊人如果要搞测量，总要在地上摊开一个全尺寸的图样，连比例概念都给忘了。如果盎格鲁-撒克逊人要绘制一张英格兰地图，那这张地图就得跟英格兰一样大。他们也用传统尺寸绘制过地图，但并不太精确。

直到 15 世纪末叶，数学活动的中心才重新回到欧洲。在中东和远东的创造力枯竭的时候，欧洲再次焕发了生机，开始从罗马教廷及其对新事物的威慑之下挣脱出来。具有讽刺性的是，这次的知识运动的中心就是罗马教廷起火的后院——意大利。

这项摧枯拉朽的变革是从一本出版物开始的。1202 年，比萨的莱昂纳多（Leonardo of Pisa）写了一本《算盘书》（*Liber Abbaci*），很久以后，人们昵称他为费波那契（Fibonacci），意为波那乔（Bonaccio）的儿子，虽然这个名字是人们在 19 世纪才创造出来的，但他就是以这个名字而闻名于世的。莱昂纳多的父亲吉莱尔莫（Guilielmo）是布嘉（Bugia，即现在的阿尔及利亚）的一位海关官员，肯定与来自各

种文化的人打过交道。他向儿子传授印度人和阿拉伯人发明的新奇数字符号，这些符号就是我们现在使用的十进制的从0到9的数字符号的前身。后来，莱昂纳多写道："在前往埃及、叙利亚、希腊、西西里和普罗旺斯的旅途中，我从不断的数学研究中获得了极大的乐趣，并在与当地学者的辩论中获益良多。"

乍看标题，我们会觉得这是一本关于算盘的书。算盘是一种运算装置，通过在金属线上滑动的珠子，或者沙制凹槽中的鹅卵石进行运算。但是就像拉丁语中的"calculus"起初表示这些鹅卵石中的一块，后来却获得了不同且更专业化的意义一样，"abbaco"一词也不再专指计算设备，而获得了"计算之艺术"的意义。《算盘书》是将印度—阿拉伯符号和方法引进欧洲的第一个算术文本。这本书的很大一部分把算术应用到了实际问题中，比如货币兑换。

一个关于野兔在理想状态下的繁殖数量问题产生了下面这个独特的数列：1，1，2，3，5，8，13，21，34，55，如此等等，每前面的两个数字之和都正好等于后面的数字。"费波那契数列"是他最著名的贡献。这一发现之所以让他声名远播，并不在于它对野兔生育问题的意义（它在这一问题上其实毫无意义），而是因为这一独特的数学形式本身及其在无理数理论中的重要作用。莱昂纳多自己都不知道这小小的数列足以使他一生中所有的其他工作都黯然失色。

莱昂纳多还写过其他一些书，他1220年的《实用几何》（*Practica Geometriae*）继承了很大一部分欧几里得几何学以及一些古希腊的三角学。欧几里得《几何原本》的第十卷讨论了由嵌套式平方根组成的无理数，其形式为$\sqrt{a+\sqrt{b}}$。莱昂纳多证明了这种形式的无理数对解三次方程是不够的。这并不是说三次方程的平方根不能通过尺规作图法作出，而是因为平方根的其他形式的组合是可以得出结果的。这第

一次表明了三次方程只用欧几里得的方法是不能求解的。

1494年，卢卡·帕乔利（Luca Pacioli）将既有的算术、几何与比例方面的所有数学知识汇总在一本书中，其中包括印度—阿拉伯数字、商业数学、欧几里得著作概要以及托勒密的三角学。这本书一个贯穿始终的主题就是大自然的设计要素都可以具象化于各种比例之中，比如人体、艺术中的透视法以及色彩的原理。

帕乔利继承了“修辞”代数的传统，用词句而不是用符号进行表述。他用“thing”表示未知数，在意大利语中是“cosa”，有一段时间，研究代数的人就被称为“cossist”。他还继承使用了由丢番图首创的规范的缩写法，但是没能发展它。莫里斯·克雷恩（Morris Kline）在他的杰作《从古至今的数学思想》（*Mathematic Thought from Ancient to Modern Times*）中作出了有力论述：“帕乔利的著作中几乎没有说出什么比莱昂纳多的《算盘书》更多的内容，这意味深长地揭示了1200年到1500年之间算术和代数的发展状况。事实上，算术和代数……还是建立在莱昂纳多那本书的基础上。”

帕乔利在其著作的末尾认为，解三次方程并不是比求一与已知圆面积相等的正方形高深多少的问题。但情况很快就变了，第一个重大突破发生在16世纪前30年间的博洛尼亚，起初并未引起人们的注意。

*

吉罗拉莫·卡尔达诺是米兰律师法吉奥·卡尔达诺（Fazio Cardauo）和一个名叫齐亚拉·米奇利亚（Chiara Micheria）的寡妇的私生子，此前米奇利亚和前任丈夫已经生有三个孩子。1501年，吉罗拉莫出生在属于米兰公爵领地的帕维亚镇。鼠疫降临米兰时，怀孕的齐亚拉在人

们的劝说下搬到了乡下，并在那里生下了吉罗拉莫。而她另外三个留在米兰的孩子，全部死于鼠疫。

吉罗拉莫在自传中写道：

> 我的父亲穿着紫色的斗篷，这种衣服在我们周围是不常见的；他时刻不忘戴着一顶黑色无檐帽……他从55岁起牙齿就掉光了。他熟稔欧几里得的著作，事实上，他的才能都是通过不断的学习得来的……我的母亲很易怒，记忆力极强而且聪慧，是一个丰满的、虔诚的小个子女人。父母两人都是急性子。

虽然法吉奥的职业是律师，但他还精通数学，甚至给达·芬奇提出过几何方面的建议。他在帕维亚大学和米兰省一个叫“皮亚蒂基金会”（Piatti Foundation）的机构教几何，并向他的私生子吉罗拉莫传授数学和占星学。

> 在我很小的时候，父亲就教给我基本的算术知识，同时还让我接触神秘事物，他在这些方面有很多我无法明白的知识。不久以后，他又教给我基本的阿拉伯占星学……我12岁的时候他开始给我讲欧几里得著作的前六卷内容。

孩子的健康状况不是很好。想让他继承家庭事业的尝试落空了，吉罗拉莫劝服了持怀疑态度的父亲让他到帕维亚大学学习医学，事实上他父亲更希望他学法律。

1494年，法国的查理八世（Charles Ⅷ）入侵了意大利，随后的战争断断续续地持续了50年。帕维亚大学因交战而关闭了，吉罗拉

莫转到帕多瓦（Padua）继续学业。无论从哪方面讲，吉罗拉莫都是个尖子生，法吉奥死时，他正在竞选大学校长的职位。虽然很多人不太喜欢他的直言不讳，但他还是以一票的优势获得任命。

在挥霍完父亲的遗产后，他开始赌博，在其混乱的人生余年里，他一直都没能戒掉赌瘾。而且他的古怪嗜好还不止于此。

> 从人生的很早时期开始，我就已严肃地投入到各等级剑术的练习中。最终，通过艰苦的训练，即使在最勇武的人当中，我也出类拔萃。……夜里，即使公爵明令禁止，我也要佩带着武器逡巡于我所居住的城市……我用一条黑色的羊毛头巾遮住自己的脸，穿着羊皮做的鞋子……经常在城市里整夜地四处游荡直到天明，流着汗演奏着小夜曲。

这真让人难以想像。

在1525年获得了医学学位之后，吉罗拉莫想要加入米兰的物理学家学会，但是被回绝了——名义上说是因为他是个私生子，其实，更大程度上是因为人们都觉得他太不谙世故。由于未能加入这个有名望的学会，吉罗拉莫就在萨科（Sacco）附近的村庄行医。这给他带来了微薄的收入，但是时好时坏。吉罗拉莫娶了军队统帅的女儿卢西亚·班达莉妮（Lucia Bandarini）为妻，把家搬到离米兰更近的地方，希望多挣点钱以维持家用，但是那个学会又一次将他拒之门外。由于不能合法地行医，他又开始赌博，但连他高超的数学技巧都没能让他富裕起来。

> 也许我压根就不值得拥有赞美，由于对棋盘和骰桌极度痴

迷，更多人认为我只配得到最严厉的指责。很多年来我既玩象棋又玩骰子，象棋四十多年，骰子差不多二十五年。我并不是每年都赌——这么说我自己都挺羞愧——而是每天都赌，并且会立刻失去想法，失去物质，还有时间。

最后他家徒四壁，家具和卢西亚的珠宝早就被当掉了。“我的人生本是光明而体面的，但是我远离了名誉和收入，却拥抱着无谓的炫耀和不合宜的快乐！我害了我自己！我走向了毁灭！”

他的第一个孩子出世了。

在经历了两次流产以及两个只活了四个月的男婴之后，以至于我……时常怀疑是由于一些邪恶的影响，我的妻子产下了我的第一个儿子……他的右耳失聪……左脚只有两个脚趾……被一块肉膜连在一起。他的背有点驼，但未到畸形的程度。这个男孩平静地生活了23年，后来，他恋爱了……娶了一个没有嫁妆的妻子，名叫布兰多尼亚·迪·瑟罗尼（Brandonia di Seroni）。

这时，吉罗拉莫过世的父亲七绕八拐地帮上了他们的忙。法吉奥在大学的授课职位一直空缺着，吉罗拉莫得到了这个工作。另外，尽管没有执照，他还间或做一些医生的工作。几个起死回生的病例——从那时的医疗状况来看，可能是由于幸运——给他带来了很高的名望。甚至那个物理学学会的有些成员也拿着自己的医学问题向他请教，在那段时间，看起来他蛮有可能进入那个受人敬重的机构。但是后来，天生多嘴的吉罗拉莫发表了一篇文章，尖刻抨击了该机构成员的才能和素质。吉罗拉莫很清楚自己的不谙世故，但并不认为这是

什么缺陷："作为一名讲师和辩论家，我既真诚又准确，只是不够精明。"1537年，由于不够精明，他的最后一次申请也被拒绝了。

但是他的名声越来越大，以至于最后该学会没有了抉择的余地，两年之后他就变成了其中的成员。吉罗拉莫的情况开始好转，而在他出版了两本数学方面的著作之后，就更加顺风顺水了。吉罗拉莫的事业开始在好几个方面齐头并进。

*

就在这段时间前后，塔尔塔里亚取得了一个重大发现——他解决了很多种类的三次方程。几经劝说，塔尔塔里亚才很不情愿地把自己的伟大发现告诉了卡尔达诺。六年之后，在塔尔塔里亚拿到一本卡尔达诺的代数著作《伟大的艺术，或代数的法则》（*The Great Art, or on the Rules of Algebra*）的抄本时勃然大怒就不足为怪了，因为这本书就是以他的发现为基础而展开的。

卡尔达诺并未剽窃荣誉，因为他对塔尔塔里亚满怀谢意：

> 如今，博洛尼亚的希皮奥内·德尔·费罗（Scipione del Ferro）解决了三次方和一次方之和等于一个常数的问题，这是一项非常了不起的成果……可与之相提并论的是我的朋友，布雷西亚的尼柯洛·塔尔塔里亚……他在与费罗的学生安东尼奥·玛利亚·费奥（Antonio Maria Fior）的论战中解决了同样的问题，并在我的百般恳求之下，把它教给了我。

但是，塔尔塔里亚看到自己珍贵的秘密被公诸于世却焦虑起来，更

让他焦虑的是，人们只会记住这本书的作者而不是这个秘密的发现者。

至少，大部分关于这件事的现存证据都说明塔尔塔里亚对此事抱有这种看法。就像理查德·维特莫（Richard Witmer）在自己翻译的《伟大的艺术》中所指出的："反仅依靠塔尔塔里亚的书面描述，很难想像有什么客观性。"后来，卡尔达诺的仆人拉多维可·费拉里（Lodovico Ferrari）称，卡尔达诺和塔尔塔里亚在那次会面中并没有关于保守秘密的协议。费拉里后来成了卡尔达诺的学生，他解决了（或者在别人的帮助下解决了）四次方程的问题，所以并不能认为他比塔尔塔里亚更客观。

对可怜的塔尔塔里亚来说，更糟糕的并不是他失去了荣誉。在文艺复兴时期的欧洲，数学发现就意味着实实在在的现金，不只是通过卡尔达诺所喜爱的赌博方式，还通过公开竞赛。

人们常说，数学不是一种面向观众的活动，但是在16世纪却非如此。那时候，数学家以公开的相互挑战为生是很正常的。在论战中，每个人都可以给对手设置一系列问题，谁的正确答案多，谁就是赢家。这虽比不上徒手搏斗和比剑扣人心弦，但即使那些旁观者根本不知道这些论战是怎么回事，也照样会出钱打赌哪一方会获胜。除了论战奖金，获胜者还会吸引一批学生，这些学生都是要交学费的。因此，公开竞赛实在是一举两得。

*

塔尔塔里亚并不是第一个发现三次方程的解法的人。1515年前后，博洛尼亚教授费罗就解出了若干类型的三次方程。费罗死于1526年，他的论文和教职都由他的女婿安尼鲍尔·德尔·内佛（Annibale Del

Nave）继承了。在 E. 巴尔托洛蒂（E. Bartolotti）的努力下，这些论文在 1970 年前后公开于博洛尼亚大学图书馆，我们这才得以确知这些事情。巴尔托洛蒂认为，费罗可能发现了三种解三次方程的方法，但是他只传下来其中一种：三次方加某项等于一个数。

内佛和费罗的学生安东尼奥 · 玛利亚 · 费奥把这种解法保存了下来。正是费奥决定以一名教师的身份用这种方法来捞钱，这家伙很有生意头脑。1535 年，他挑起了与塔尔塔里亚关于三次方程问题的论战。

传闻说有人已经找到了三次方程的解法。没有什么能比找到解决问题的方法更能鼓舞数学家的了。人们都不愿意在没有答案的问题上白费工夫，最危险的是，一个问题确实是有答案的但你却没能找到。你只需要强大的自信心就够了，数学家是很少缺乏自信的，就算有时只是盲目自信。

塔尔塔里亚也发现了费罗的方法，但是他怀疑费奥还知道一些其他类型的三次方程的解法，因而还保有很大的优势。所以，塔尔塔里亚在论战之前的很短时间内解决了其余类型的三次方程，其担忧可见一斑。最后塔尔塔里亚一举击败了倒霉的费奥。

这场论战的消息不胫而走，卡尔达诺在米兰听说了此事，当时他正在写自己的代数著作。像所有真正的作者一样，卡尔达诺决心把最新的发现都写进著作里去，因为如果没有最新发现，这本书在出版之前就已经过时了。于是卡尔达诺找到了塔尔塔里亚，希望能把他的发现套出来，以便写进自己的《伟大的艺术》中。但是塔尔塔里亚拒绝了，他说他自己也打算写本书。

但是最终，由于卡尔达诺坚持不懈的努力，塔尔塔里亚透露了自己的秘密。在他得知卡尔达诺那本书即将面世的情况下，到底有没有要卡尔达诺发誓保守秘密呢？或者是他屈从了卡尔达诺的哄骗但是之

后又后悔了？

无疑，在《伟大的艺术》面世之后他是极度愤怒的。没过一年，他就出版了自己的《问题和发明种种》（*Diverse Questions and Inventions*），在书中毫不含糊地攻击卡尔达诺。他详细精确地列出了两本书之间的全部相似之处。

1574年，费拉里站出来支持他的主人。他发起了一场学术辩论，内容包括塔尔塔里亚涉及的所有主题。他甚至给出了200斯古蒂（scudi，意大利16~19世纪货币单位）的奖金。费拉里的观点很明确："我要让人们明白，你写的东西都是不实且可耻的诽谤……与西格诺·吉罗拉莫相比，你简直就不值一提。"

费拉里将辩论公告分发给了意大利很多学者和公众人物。九天后，塔尔塔里亚以他自己对事实的陈述作出回应。两位数学家之间的公告往来了12个回合，时间达18个月之久。这场辩论看起来是遵循着真正的决斗的标准规则进行的。受尽费拉里侮辱的塔尔塔里亚被允许挑选自己的武器——选择一个辩论主题。但是他坚持要和卡尔达诺论辩，而不是和他的挑战者费拉里。

费拉里平心静气地指出，从任何方面讲，最先解出三次方程的都是费罗，而不是塔尔塔里亚。既然费罗都没有因塔尔塔里亚对荣誉的不正当要求而大呼小叫，为什么塔尔塔里亚就不能这样呢？这一论点非常有力，塔尔塔里亚也许意识到了这一点，因而打算退出论战。但是他没有这么做，这可能是因为其故乡布雷西亚的几个市参议员。塔尔塔里亚正在那里谋求一个讲师职位，这些达官贵人们可能想看看他如何为自己脱罪。

不管怎样，塔尔塔里亚还是答应了这场辩论。辩论于1548年8月在大庭广众之下于米兰一座教堂内进行。除了塔尔塔里亚的只言片

语，并没有关于辩论过程的记录。塔尔塔里亚说，辩论结束于晚饭之前。这暗示辩论可能并不太扣人心弦，但是看起来费拉里轻而易举地取得了胜利，因为之后他获得了一个肥缺，成为米兰总督的税务审核员，不久就腰缠万贯了。另一方面，塔尔塔里亚从没说过自己赢了这场辩论，也没有得到布雷西亚的那份工作，并且陷入了痛苦的反诉之中。

塔尔塔里亚不知道，卡尔达诺和费拉里有一套与他截然不同的辩论策略。他们走访了博洛尼亚大学，并且翻阅了费罗的论文，包括那个最初提出的三次方程的解法。以后的几年中，他们两人都坚称《伟大的艺术》中的内容都来自费罗的原始著作，而不是塔尔塔里亚教给卡尔达诺的，之所以提到塔尔塔里亚的发现只是为了说明卡尔达诺自己是如何得知费罗的著作的。

但故事最后峰回路转。1570 年，就在《伟大的艺术》发行第二版之后，卡尔达诺被宗教裁判所监禁了。监禁理由完全是无稽之谈。不是因为这本书的内容，而是因为其献辞。卡尔达诺把这本书献给了一个名不见经传的知识分子安德里亚斯·阿西安德尔（Andreas Osiander），一个宗教改革中的小人物。但是，人们很怀疑这个人就是哥白尼《天体运行论》的匿名前言的作者。《天体运行论》是第一本提出行星围绕太阳，而不是地球运行的书。教会把这种观点视为异端，并在 1600 年活活烧死了坚持这一观点的布鲁诺——在罗马一处市场的广场上将其赤身裸体地倒挂在火刑柱上，还堵上了他的嘴巴。1616 年和 1633 年，同样的原因给伽利略带来了巨大的痛苦，但是那时宗教裁判所已能满足于仅仅是将他软禁起来。

*

要理解吉罗拉莫及其同胞们的成就，我们必须回去看看解释二次方程解法的巴比伦泥板。如果按照他们的方法并用现代符号来表达其运算步骤，我们会发现，巴比伦人对二次方程 $x^2-ax=b$ 的解法是：

$$x=\sqrt{\left(\frac{a}{2}\right)^2+b}+\frac{a}{2}$$

这就是当年每个小学生都要熟记的公式，在当今的公式论著中已经随处可见了。

文艺复兴时期的三次方程解法与此相似，但是更精密一些。用现代符号表示，比如说 $x^3+ax=b$，那么

$$x=\sqrt[3]{\frac{b}{2}+\sqrt{\frac{a^3}{27}+\frac{b^2}{4}}}+\sqrt[3]{\frac{b}{2}-\sqrt{\frac{a^3}{27}+\frac{b^2}{4}}}$$

在这类公式中，这一个还是相对简单的（相信我！），但是想要描述它，你非得事先学习大量的代数概念不可。到目前为止，这是我们将会遇到的最复杂的公式了，它用到了我介绍过的全部三种符号：字母、上标的数字和$\sqrt{\ }$这个符号，它既表示平方根也表示立方根。你没必要理解这个方程，当然也没有必要对它进行演算。但是你得理解它的一般形式。首先，有些术语在我们接下来的论述中是非常有用的。

所有像 $2x^4-7x^3-4x^2+9$ 之类的代数表达式都叫“polynominal”，意思是“多项式”。这样的代数式是由未知数的不同次幂相加构成的。与未知数的幂相乘的数字 2，−7，−4 和 9 叫做系数。方程中未知数的最高次幂叫做多项式的“次数”，所以这个多项式的次数就

是4。低次（从1次到6次）多项式都有自己的专有名称：线性方程（linear）、二次多项式（quadratic）、三次多项式（cubic）、四次多项式（quartic）、五次多项式（quintic）和六次多项式（sextic）。组合方程$2x^4-7x^3-4x^2+9=0$的解叫做多项式的根。

我们现在来分析一下卡尔达诺的公式。它由系数a和b构成，使用了加减乘除法（但是只能被整数除，即2，4和27）。这个公式有两个深奥之处：一个是平方根，事实上同一个平方根出现在了两个不同的地方，但一个是加一个是减；最后是两个立方根，而且是含有平方根的立方根。因此，除了无伤大雅的代数处理（我是指单纯地来回移动各项）外，这种解法就可以概括为："开平方，然后开立方；重复做一次，然后把两次的得数加起来。"

做这么多就够了，但是我认为再少可就不行了。

文艺复兴时期的数学家们没有发现，这并不只是某一类型的三次方程的解法，但后来者很快就意识到了这一点。这种解法适用于所有类型的三次方程，只要作一些简单的代数处理就可以了。如果三次方项是$5x^3$而不是x^3，你可以将整个方程除以5——这一点文艺复兴时期的聪明数学家肯定能想到。但是一种更加微妙的思想则需要一种对数的革命性见解，那就是在必要时允许系数a和b是负数，这样就可以不用再费力区分各种情况下的方程了。最后就是纯粹的代数技巧了：如果方程含有未知数的平方，你可以把它消除——你可以把x变成x加上一个悉心选定的常数，如果你选对了，平方项就会神奇地消失。文艺复兴时期的数学家们担心那一项会消失无踪，但是现在看来，这一项并不是消失了，只不过是系数变成了零，方程还是同一个方程。

问题解决了吗?

没有彻底解决。我说谎了。

我的谎言在于：我说卡尔达诺解决了所有类型的三次方程。但是在某种意义上，这么说是不对的，事实证明，这一点非常重要。但是我谎撒得并不离谱，这就看你怎么理解“解决”一词的含义了。

卡尔达诺自己也意识到了其中的困难，这一困难很好地解释了他对待细节的态度。三次方程一般有三个解（排除一个负数解，就没有这么多了）或一个解。卡尔达诺发现，当一个方程有三个解时——我们可以将它们命名为1，2，3——这个公式并不全然适用于这三个解。因为，公式中可能包含有一个负数的平方根。

这里有一个明显的例子。卡尔达诺发现三次方程 $x^3=15x+4$ 的一个很明显的解是 $x=4$。但是，当他用这个结果检验塔尔塔里亚的公式时，得出的答案是 $x=\sqrt[3]{2+\sqrt{-121}}+\sqrt[3]{2-\sqrt{-121}}$，这看起来根本就讲不通。

在当时的欧洲数学家当中，几乎没人有足够的勇气去考虑负数，而他们的东方同道却很早就开始思考负数问题了。早在公元5世纪，印度的耆那教徒就提出了最基本的负数观念；13世纪，中国的“筹算”体系就用红筹表示正数，黑筹表示负数——虽然只存在于非常有限的文本中。

如果说负数是一个难题，那它的平方根就更令人头疼了。问题的难点在于，无论正数还是负数的平方都只能是正数——我现在不会去对其进行解释，但是我必须说明，只有这样，代数规则才能协调一致地发挥作用。因此，无论你多么迷恋负数，也必须承认它没有合理的平方根。所有包含负数的平方根的代数式都是无意义的。

但是，卡尔达诺根据塔尔塔里亚的公式得出的正是一个这样的代

数式。如果你知道在某种情况下一个方程有一个可以通过其他途径求出的解，但却无法通过这一公式求出，确实很让人不安。

1539年，忧心忡忡的卡尔达诺向塔尔塔里亚提出了疑问：

> 我已经认真研究了很多你给我的那些不带答案的问题的解法，其中之一是三次方等于未知数加上一个数，我已经牢牢掌握了这种问题的法则。但是当未知数的系数的三分之一的三次方值大于这个数的二分之一的平方时，我发现它就不适合这个方程了。

卡尔达诺此处描述的正是负数的平方根问题。很明显，卡尔达诺对这一问题的整体理解非常了不起，并且发现了一个大难题。我们并不清楚塔尔塔里亚对自己的公式的理解是不是也达到了同样的水平，因为他的回答是这样的："你还没有掌握解决此类问题的正确方法……你的方法是完全错误的。"

可能是塔尔塔里亚只是存心不愿意帮忙，也可能是他没有明白卡尔达诺所指的到底是什么。不管怎么说，卡尔达诺已经涉足了困扰全世界数学家250年之久的难题。

*

其实在文艺复兴时期，已经出现了某种重大变革的迹象。《伟大的艺术》中讨论了由同一议题引起的另一问题：找出和等于10，积等于40的两个数。得出的答案是$5+\sqrt{-15}$和$5-\sqrt{-15}$。如果你忽略-15的平方根意味着什么（指其合理性）的问题，就当它和正常的平方根一样，就会发现这两个数正好适合这个方程。如果把这两个数相加，

两个平方根就抵消了，而两个 5 相加等于 10，正好符合要求；如果这两个数相乘，就会得到 $5^2-(\sqrt{-15})^2$，等于 25+15，正好是 40。卡尔达诺不明白为什么会出现这样古怪的方程。“因此，”他写道，“算术技巧推进到越精致的地步，就越无用。”

一个羊毛商人的儿子拉费利·邦贝利（Lafaele Bombelli），在自己 1572 年出版的《代数学》（*Algebra*）中提到了类似的运算。把一个“想像的”平方根当成一个真实的数字乘以自己，就能把卡尔达诺令人费解的三次方问题的古怪公式变换成正确答案：$x=4$。他写这本书是为了打发自己的业余时间，当时他正在为教廷的法律和财政部门（教廷办事处）收回一片沼泽地。邦贝利发现，$(2+\sqrt{-1})^3=2+\sqrt{-121}$，还发现 $(2-\sqrt{-1})^3=2-\sqrt{-121}$。因此，这两个立方根的和就变成了 $(2+\sqrt{-1})+(2-\sqrt{-1})$，正好等于 4。无意义的平方根不知何故变得有意义了，并且得出了正确答案。邦贝利也许是第一位意识到带有负数的平方根的代数运算可以实现，并能得出合理结果的数学家。这强烈地表明了这些数存在着一个合理的解释，但他并没有指出该对其作何解释。

*

卡尔达诺的著作的最高数学成就并不是三次方程，而是四次方程。他的学生费拉里发展了塔尔塔里亚和费罗含有未知数的四次方的方程的解法。费拉里的公式只牵涉到了平方根和立方根——四次方根只是平方根的平方根，所以并不是必须的。

《伟大的艺术》并没有涉及含有未知数的五次方的五次方程。但是随着方程次数的增加，其解法也随之越来越复杂了，几乎没有人怀

疑，只要有足够的天才，五次方程应该也是可以解决的——你可能不得不用到五次方根，而且公式会极为复杂。

卡尔达诺并没有下工夫去探求这一问题的解法。1539 年之后，他又回到了他的其他活动中，尤其是医学。这时，他的家庭以最可怕的方式分崩离析了："我的儿子（最小的一个）结婚不久就受到了指控，说他毒死了正处于分娩期的虚弱的妻子。2 月 17 日，他被逮捕了，53 天之后，也就是 4 月 13 日，他在监狱被砍了头。"在卡尔达诺努力与苦难达成妥协的时候，却祸不单行。"一座房子——我的房子——在短短的数日之内经历了三次葬礼，儿子的、小孙女蒂亚雷吉娜的和孩子的奶妈的，我那襁褓中的孙子也快要一命呜呼。"

尽管如此，对于人类的状况，卡尔达诺却是不可救药的乐观主义者："但是，我仍然有着这么多幸事，如果这些事降临在别人头上，那他也应该觉得自己是幸运的。"

狡狐

到底该走哪条路？研究哪一个学科更好？必须在自己同样喜欢的两个学科之间作出选择，这确实让人左右为难。1796 年，一个才华横溢的 19 岁青年面临着将要影响自己一生的抉择，他不得不确定自己的职业。虽然来自一个普通家庭，但是卡尔·弗里德里希·高斯（Carl Friedrich Gauss）明白自己会成为一个了不起的人。包括布伦瑞克公国（Brunswick）的公爵在内，大家都很赏识他的才华。高斯就出生在公爵的领地上，他一家人也都生活于此。问题是高斯在很多方面都很有才华，而且必须在自己最喜欢的两个科目——数学和语言之间作出选择。

但是，3 月 30 日，这一抉择上的难题被一个神奇、非凡并且史无前例的发现解决了。这一天，高斯用欧几里得作图法作出了正十七边形。

这看起来让人难以理解，因为在欧几里得的著作中，没有关于这一问题的任何线索。在他的著作中，你可以找到 3 条边、4 条边、5 条边或 6 条边的正多边形的作法；也可以通过三边形和五边形作图法

的混用作出十五边形；还可以通过反复地等分使边数整倍增加，以此可得到边数为 8，10，12，16，20……的正多边形。

十七边形让人很头疼。但它很典型，高斯对它的典型性了如指掌。这可以归因于数字 17 的两个基本性质：它是一个质数，只能被自己和 1 整除；它还比 2 的一个幂大 1，$17=16+1=2^4+1$。

如果你是高斯那样的天才，就会发现这两个普通的命题暗含着一种图形的作法，这就是用直尺和圆规作正十七边形；如果你是生活在公元前 500 年到公元 1796 年之间的另一位伟大数学家，恐怕连这种联系的蛛丝马迹都找不到一丁点儿。我之所以这样说，是因为确实没人做到。

借此，高斯就足以验证自己的数学天赋了。因此，他决定做一位数学家。

*

1740 年，高斯的祖父在布伦瑞克谋得了一个园丁的职位，他们举家搬到了那里。三个儿子当中，格布哈德·迪特里希·高斯（Gebhard Dietrich Gaush）也成了一名园丁，也不时干些码砖、修整沟渠之类的体力活，另外，他还是“水利师傅”、商务助理和一项保险基金的出纳员。高利润的交易活动都在行会的手里，初来的移民，甚至移民的后代都被排斥在外。1776 年，格布哈德娶了第二个妻子多萝西娅·本茨，她是个石匠的女儿，做些女仆的活计。他们的儿子约翰·弗里德里希·卡尔（他们总是叫他弗里德里希·卡尔）出生于 1777 年。

格布哈德很诚实，然而固执、粗暴，而且不大机灵。多萝西娅精明而有主见，这些特点深深影响了高斯。孩子两岁的时候，多萝西娅

发现他是一个神童，她倾尽心力要让孩子接受教育，以使他的天分得到充分发展。格布哈德更希望卡尔做一个瓦匠。儿子19岁时，多萝西娅的朋友沃尔夫冈·波尔约（Wolfgang Bolyai）预言卡尔将成为“欧洲最伟大的数学家”，多亏了这位母亲，卡尔把预言变成了现实。听见这个预言时，多萝西娅竟然喜极而泣。

这个孩子报答了母亲的奉献，多萝西娅生命中的最后20年，和他生活在一起，这段时间，她的视力一直在下降，最后变成了瞎子。这位杰出的数学家一直照料着自己的母亲，他看护着她，直到她1839年去世。

高斯很早就展露了自己的天才。3岁的时候，他看父亲（他当时是个工头）计算这个月的工资，发现了一个计算错误并指了出来，这让格布哈德大吃一惊。

几年之后，一位名叫J. G. 布特内尔的教师给全班同学布置了一项作业，这项作业会让学生们忙上几个钟头，好让老师好好休息一下。我们并不太清楚这个问题具体是什么，但大致是下面这种问题：把从1到100的数全部加起来。当时很可能并不是这么简单的数字，但肯定是这样的潜在规律：有一定的数学连续性，每两个相邻数字的差值都是相同的。这些数字有一个简单但是并不十分明显的计算技巧，但学生们都还没有学过，因此就必须辛苦地把这些数字逐个加起来。

至少，布特内尔就是这么想的。他要求学生们尽快完成自己的作业，把答案写在石板上，放在自己的桌子上。于是，他的这群学生开始涂抹起来：1+2=3，3+3=6，6+4=10。免不了还会出10+5=14之类的错误。高斯没有跟着大家涂写，而是想了一会，在自己的石板上写下了一个数字，把石板面朝下放在桌上，向老师走去。

“我已经写下来了。”说完，他回到桌边坐下了。

这堂课结束的时候，老师把所有的石板收上来，只有一个上面写着正确答案——高斯的。

我们仍然不知道高斯的思维过程，但可以提出一种可信的设想。高斯很有可能已经思考过这类加法的有效运算技巧（若非如此，就证明他有立时发明一个技巧的能力）。解决这一问题的简单方法是将数字两两配对组合：1 和 100，2 和 99，3 和 98，以此类推，直到 50 和 51，全部数字的和就等于这些成对数字的和的和，正好 50 对。所以，这些数字的总合是 50×101=5050。这就是他写在石板上的（或者类似这样的）得数。

我并不是要强调高斯的数学学得很好，但事实就是如此；在他后来的天文学研究中，他按部就班地进行了大量高数位的运算，具有白痴天才一样的运算速度。但是娴熟的运算并不是他的惟一才能。他的最突出天赋是发现数学问题的隐秘形式，并用它们找到解决问题的方法。

布特内尔非常吃惊，高斯竟然识破了自己的狡计。他因此感到非常光荣，给这个男孩买了所有用钱可以买到的、最好的数学教科书。不到一个星期，老师已经没什么可以教给高斯了。

碰巧，布特内尔有一个 17 岁的助教，名叫约翰·巴特尔斯（Johann Bartels），他的工作是削尖羽茎，并教孩子们如何使用。在工作之外，他还是一个十足的数学迷。他跟 10 岁的天才男孩交上了朋友，两个人变成了终身好友。他们一起钻研数学问题，并互相激励。

巴特尔斯和布伦瑞克的一些头面人物很熟。这些人很快发现，自己身边有一位不为人知的天才，而他的家庭却生活却处在贫困的边缘。这些士绅当中的一位，议员兼教授 E. A. W. 齐默尔曼（E. A. W. Zimmerman），在 1791 年把高斯引荐给了布伦瑞克的公爵卡尔·威廉·斐迪南德（Carl Wilhelm Ferdinand）。公爵对高斯非常喜爱，决定

支付高斯的教育费用。他经常资助穷人家有天赋的孩子上学。

数学不是这个男孩的惟一天赋。15 岁的时候，高斯就熟练掌握了古典语言，因此，公爵就给地方体育馆（在从前德国的教育系统中，体育馆是学生们进行大学入学训练的学校，可以大致翻译成“高中”，但只供济获得认可的学生）出资研究古典语言。高斯的很多最好的著作都是用拉丁文写成的。1792 年，同样是在公爵的资助下，高斯进入了布伦瑞克的卡罗琳学院（Collegium Carolinium）。

17 岁的时候，高斯已经发现了一个惊人的定理，就是数论中人们熟知的“二次互反律”（Law of Quadratic Reciprocity）。这是关于完全平方的可除性的一个基础然而费解的规律。高斯并不知道欧拉已经发现了这种规律，独力作出了这一发现。很多人都没有考虑过这个问题，但高斯对这一问题的思考却十分深入。事实上，正是这一发现使他作出了正十七边形，并把他引上了成为不朽数学家的道路。

*

1795 年到 1798 年间，在斐迪南德的资助下，高斯在哥廷根攻读大学学位。他很少交朋友，但他的朋友都是密切而持久的。在哥廷根大学，高斯遇见了波尔约，他当时已是一位颇有成就的欧几里得几何学家了。

高斯的数学思想来得又猛烈又迅速，有时几乎让自己有点承受不了。当新的想法袭来时，他会突然停下手上的事情，在思维转换的间隙失神发愣。他曾经得出过一些不符合欧几里得几何的定理。他的主要思考成果集中在《算术研究》（*Disquisitiones Arthmeticae*）这部巨著中。这本书 1798 年就差不多完成了，但是他想得到前辈的认可。于

是他来到了黑尔姆施泰特大学（University of Helmstedt），那里有一座一流的数学图书馆，由德国最著名的数学家普法夫（Johann Pfaff）掌管。

1801 年，由于印刷商的原因，《算术研究》令人沮丧地被一再拖延，但最终还是出版了。高斯把这本书题献给了斐迪南德公爵，献词写得诚挚而热情洋溢。公爵的慷慨资助一直延续到高斯离开大学之后。他资助了高斯的博士论文，出钱给黑尔姆施泰特大学，将这篇论文按照正规标准印刷出来。在高斯担心自己离开大学之后如何维生的时候，公爵给他以保障，他可以安心进行自己的研究，而不必为钱的问题发愁。

《算术研究》的一个鲜明特征是它较真的写法。书中的证明严密而又条理清晰，对读者毫不迁就，而且，这本书剔除了定理背后所有直觉方面的因素。后来，他确定了自己的这一态度，并使之贯穿了尔后的整个著述生涯。这一态度的原则是："建筑已经完工，脚手架就没用了。"让人们站在建筑面前赞叹就好了，没有必要教会人们它是如何建成的。在以高斯的思想为基础进行复杂分析的著作中，卡尔·古茨塔夫·雅科布·雅科比（Carl Gustav Jacob Jacobi）谈到这位前辈时说："他就像狐狸一样，边走边用尾巴在沙地上扫除行迹。"

*

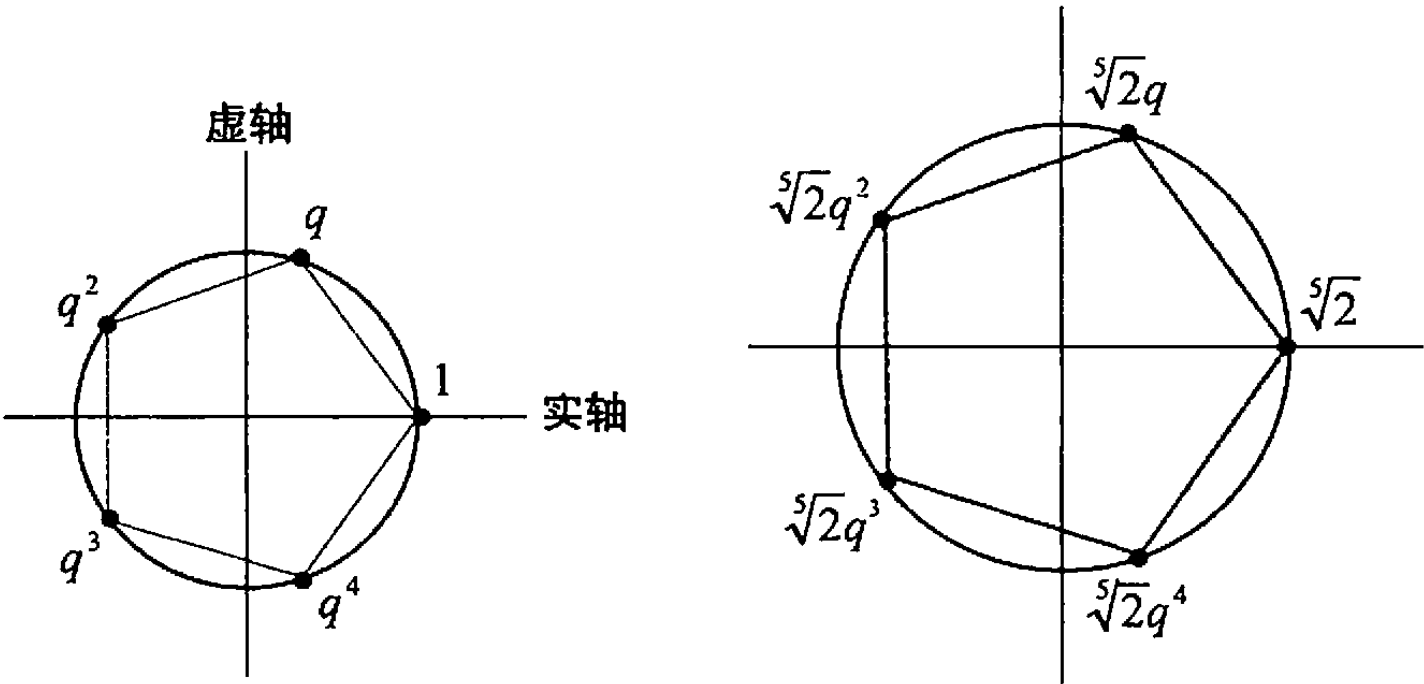
同一时期的数学家们开始逐渐意识到，尽管复数看起来复杂而费解，但却提供了一种规范的解方程方法，使代数运算更加简捷。简洁是数学的试金石，无论一些虚构观念一开始看起来多么古怪，但长远来看，只要它能让特定的问题变得简洁，都会留存下来。

如果你单靠实数来解方程，很可能会麻烦得让你气急败坏。方程

$x^2-2=0$ 有两个解，分别是 2 的正负平方根，而与它看起来差不多的 $x^2+1=0$ 却是无解的。但这个方程确有两个复数解：i 和 $-i$。表示$\sqrt{-1}$的 i 是欧拉于 1777 年引入使用的，但是这篇文章直到 1794 年才发表。单纯用实数表达的方程抛出了很多的特例和成规，但类似的复数方程通过一个较大的复杂性——既接受实数也接受复数——囊括了诸多复杂性。

到 1750 年，文艺复兴时期的数学家发起的数学思想进程成熟了。他们的三次方程和四次方程的解法看起来就像巴比伦时期二次方程解法的自然延伸。通过一些技巧弄清了根式和复数之间的联系，并且，在这个体系中，一个数字有三个而不是一个立方根；有四个而不是一个四次方根；有五个而不是一个五次方根。要弄清这些根到底从哪儿来，就必须理解“单位根”的性质，“单位根”其实就是 1 的 n 次方根。这些根是复数平面上的正多边形的定点，其中有一个定点位于 1 的位置上。在此空间中，其余的根都分布在以 1 为半径、以 0 为圆心的圆周上。如下图所示，就是五次单位根的分布情况。

推而广之，由某一数的任一五次方根都可以得出另外的四个，只需将其乘以 q，q^2，q^3 和 q^4 就可以了。这些数字也都等距分布在以 0 为圆心的圆周上。比如 2 的五次方根的分布，如右图所示。



复数平面上的五次方根；2 的五次方根

这很漂亮，而且颇具深意。2的五次方根可以看做方程$x^5=2$的解。这是一个五次方程，有五个复数解，但只有一个是实数。以此类推，方程$x^4=2$有四个2的四次方根的解，2的17次方方程有17个解，如此等等。即使你不是个方程天才，也能够想到——方程的解的数量就等于它的次数。

这并不仅仅适用于求n次方根的方程，而且适用于任何代数方程。数学家们逐渐发现，在复数领域内，每个方程的解的数量都正好等于方程的次数（确切说来，这一命题只有按照这些根的多重性来计算其个数时才能成立，没有这个限定条件，方程解的个数就会少于或等于方程的次数）。欧拉证明了二次、三次和四次方程的这种性质，并断言了这种方法的普适性。他的思想是令人信服的，但是充满了非常严重的缺陷，甚至直到今天，人们还在为把他的方法变成一个结论而大费脑筋。尽管如此，数学家们却认为，既然要解方程，就要精确地求出这一方程的每一个解。

在提出了自己的数论思想并作出分析之后，高斯对这一没人能够证明的假设越来越不满。按照高斯一贯的行事风格，他提出了一项证明。这一证明曲折复杂：高斯的竞争对手们只有表示信服的份儿，而不知道高斯得出这一证明的切入点。这只狐狸又把行迹扫除了。

*

高斯的拉丁文论文的标题可以翻译成“关于每一个变量的有理整数函数都可以分解成一次或二次方程的实数因数的新证明”。高斯的论文并没有纠缠于当时的术语，正如其题目所揭示的，每一个（系数为实数的）多项式都是线性方程或二次方程的组合。

高斯以“实数”一词表明自己是在传统的数论体系中展开研究的，在这一体系中，负数是没有平方根的。今天我们可以条理清晰地陈述高斯的这一定理：每一个 n 次多项式都有 n 个实数或复数解。但是高斯措辞十分谨慎，所以他并没有借助复杂的复数体系。实数多项式的复数根总是成对出现，这些成对出现的复数根产生相当于实数根的二次实数因数、线性实数因数。高斯通过题目的断句区分出两种类型的因数（一次、二次因数），因此避开了可能引起争论的复数问题。

论文题目中有一个毫无来由的词——“新的”，借以暗示此前已经有过对这一问题的证明。高斯第一次给出了对这一基本定理的严密的代数证明，但又不愿触犯那些声称已经证明这一问题的著名的先行者——其实他们都错了——高斯称自己的突破性进展为最新的证明，只是用了一种新方法。

这一定理就是后来的代数基本定理。高斯对这一定理非常看重，以至于作出了四种证明，而最后一种证明是在他 70 岁的时候作出的。他个人对复数深信不疑，复数在他的思想中举足轻重，他还陆续不断地提出了自己对复数意义的解释。但是，高斯厌恶争论。后来，他推翻了自己很多的原创性思想，其中包括非欧几何、复数分析，以及对复数的严密研究，因为他实在不想听到他所说的“愚夫的哭闹”。

*

高斯并没有拘泥于理论数学。1801 年初，意大利传教士兼天文学家朱塞佩·皮亚奇（Giuseppe Piazzi）发现了一颗新行星，或者说是他自认为如此——他在自己的望远镜中看到了一线微弱的光芒闪耀在昼夜交替时的星群中，借此，他确信这是太阳系里的一颗星体。这一星

体被正式命名为谷神星（色雷斯，Ceres），它事实上就是第一颗被人所发现的小行星。但是很快，皮亚奇发现的这一崭新的世界消失在耀眼的太阳光里。他的观察少得可怜，以至于天文学家们无法确定这颗行星的运行轨道，他们甚至担心，这颗行星隐没到太阳背后时必须重新对其进行定位。

这对高斯来说是个有意思的问题，所以他满怀热情地投入其中。他发明了通过较小的数确定行星轨道的更佳方法，并预测了谷神星再次出现的位置。在他的预测成为现实之后，高斯名闻遐迩。勘探学家亚历山大·冯·洪堡（Alexander von Humboldt）问天体力学家皮埃尔–西蒙·德·莱普雷斯（Pierre-Simon de Laplace）谁是德国最伟大的数学家时，得到的答案是“普法夫”。亚历山大·冯·洪堡又惊奇地问道：“那么高斯呢？”莱普雷斯回答道：“高斯是全世界最伟大的数学家。”

不幸的是，这一新发现把他从纯粹数学领域拖入了漫长而乏味的天体力学运算之中——人们普遍认为这是对他的杰出天赋的浪费。我并不是说天体力学不重要，而是几乎没有哪个数学家完成过这样一项工作。另一方面，这件事也让高斯开始受生活之累。当时，高斯正在谋求一个体面的公共服务方面的差事，想以此来报答他的支持者斐迪南德公爵。但是他对谷神星的研究工作让他成了哥廷根观测站的带头人，并在那里度过了之后的学术生涯。

1805年，高斯与约翰娜·奥斯托弗（Johanna Osthoff）结婚。在给波尔约的信中，高斯这样描述他的新任妻子：“其一，圣母玛利亚一样漂亮的脸，映透着宁静的心灵，健康、温柔，灵动的眼睛，容貌无可指摘；其二，聪明的头脑和很高的语言修养；最重要的是，她安静、安详、温和，拥有天使一样纯真的灵魂，善待一切。”约翰娜为高斯生了两个孩子，但是她在1809年死于难产，悲痛欲绝的高斯“帮

她合上天使般的眼睛，在过去的五年中，我在这双眼睛中找到了天堂”。他陷入了孤独和沮丧，从此以后，他再也没有找回以往的生活。他娶了另一位妻子，约翰娜最好的朋友米娜·沃尔戴克，但是，除了又生了三个孩子之外，这段婚姻并不太幸福。高斯经常和儿子们争吵，还得教女儿们该如何做事。儿子们厌倦了，干脆离开欧洲去了美国，并在那里开枝散叶。

成为哥廷根的带头人不久，高斯又思考起了一个老问题——一种满足除平行公理之外的所有欧几里得几何学公理的新型几何学。他最后确信，一种非欧几何在逻辑上是可能的，但是他从来没有发表这一成果，因为他觉得这一观点太过激烈了。后来，他的老朋友的儿子亚诺斯·波尔约（János Bolyai）也得出了这一发现，但由于顾虑过多，高斯并没有赞扬过他的研究成果。直到后来，尼古莱·伊万诺维奇·洛巴切夫斯基（Nikolai Ivanovich Lobachevsky）再次独立发现非欧几何，高斯只是吸纳他为哥廷根学院的正式成员，仍然没有对其进行公开表扬。

数年后，随着对这种新的几何学的研究不断细化，它们被人们解释为曲面上的“最短线的”几何学。如果这个面有稳定的正曲度，比如球面，这种几何学就叫椭圆几何。如果一个面有稳定的负曲度（形似马鞍之类），这种几何就叫双曲线几何。欧几里得几何对应的是零曲度、扁平空间。非欧几何的独特之处在于其测量方式，它用不同的公式测量两点之间的距离。

高斯可能借由这些思想对曲面展开过更具有普遍性的研究。他发现了关于曲率总和的漂亮公式，并且证明，在任何坐标系中，曲率的总和都是相同的。在这一公式中，并不要求曲率是恒定的：它在曲线的各个部分之间是可变的。

高斯在中年以后转向了实际操作，这在数学家当中是一种常见的转变。他参与了好几次测绘工作，其中规模最大的是汉诺弗三角的测绘。他做了很多实地测量工作，随后又进行了数据分析。为了辅助这项工作，他发明了回光仪，一种用反射光传递信号的装置。但是当他的心力显现出衰退的迹象后，他停止了测绘工作，决定在哥廷根度过余年。

在高斯那段不幸的时期，年轻的挪威人阿贝尔给他写信谈论了无法用根式解五次方程的问题，但是高斯没有回信。因为他当时是那样的心灰意冷，所以很可能根本就没有读这封信。

1833 年前后，高斯对电磁学产生了兴趣，并与物理学家威尔海姆·韦伯（Wilhelm Weber）合著了《地磁概论》（*General Theory of Terrestrial Magnetism*）一书，并于 1839 年出版。他们还发明了电报机，以联系高斯的观察站和韦伯的物理实验室。但是电报信号经常中断，而其他的发明家已经提出了更实用的设计。后来，包括韦伯在内的六个人因不愿与汉诺弗的新国王恩斯特·奥古斯特（Ernst August）合作而被开除出哥廷根大学。这让高斯非常丧气，由于他很保守，又不愿意制造政治影响，这一切才没有使他公开树敌，甚至也没有作出为韦伯出头的举动。

1845 年，高斯撰写了关于哥廷根大学教授的遗孀的养老金问题的报告，考察了成员迅速增加可能带来的后果。他还投资了铁路和政府债券，赚了一笔小钱。

1850 年以后，由于心脏病频频发作，高斯中断了工作。在这一时期，与他有关的最重要事件是他的学生格奥尔格·伯恩哈德·黎曼（Georg Bernhard Riemann）发表了争取任教资格的论文（德国的学员系统中，获得了博士学位之后，紧接着就要提交争取任教资格的论

文）。黎曼简单地将高斯关于多维空间的著作加以概括，名之为“拓扑面”。他特别发挥了关于测量方式的观念，发明了一个计算拓扑面曲率的公式。实际上，他创造了弯曲多维空间的理论。这一思想在后来爱因斯坦的重力论中非常关键。

这时的高斯被他的医生井井有条地照料着，他写文章赞许了黎曼的论文，这篇文章发挥了很大的作用。随着健康的进一步恶化，高斯只能越来越多地待在床上，但仍然坚持写作、阅读，还要管理自己的投资。1855年初，高斯，有史以来最伟大的数学头脑，在一场睡眠中平静地死去了。

受挫的博士和多病的天才

对卡尔达诺的《伟大的艺术》的实质性突破出现在18世纪中叶。虽然文艺复兴时期的数学家可以解决三次方程和四次方程的问题，但这些方法只不过给出了一套很基础的技巧而已。这些技巧都分别适用于不同的特定条件，而没有给出一个系统化的理论。1770年前后，两位数学家——约瑟夫-路易斯·拉格朗日（Joseph-Louis Lagrange）和一个认为自己是法国人、名叫亚历山大-乔菲尔·范德蒙德（Alexandre-Théophile Vandermonde）的意大利籍数学家，最终创立了关于这一问题的理论。

范德蒙德1735年生于巴黎。他的父亲想让他做个音乐家，范德蒙德也作为一个小提琴手从事了音乐。但是在1770年，他迷上了数学。他发表的第一篇数学论文是关于多项式方程根的对称函数的。对称函数的代数式看起来像多个根的和，如果这些根可以互换，它就是不变的。对称函数可以结合正多边形来证明，如果n是10或小于10，方程$x^n-1=0$就可以通过根式解决（事实上，n为任意数都可以解决）。

法国著名分析家奥古斯丁-路易斯·柯西（Augustin-Louis Cauchy）指出，范德蒙德是第一个发现对称函数可以用来解根式方程的人。

而在拉格朗日那里，这一思想开始被拿来对付所有的代数方程。

*

拉格朗日生于意大利都灵市，受洗名为朱塞佩·洛德维科·拉格朗日亚（Giuseppe Lodovico Lagrangia）。他的家族与法国联系紧密，他的曾祖父在前往意大利服侍萨沃伊公爵以前，是一名法国骑兵军官。朱塞佩还很小的时候就以拉格朗日为姓，但总是与名字洛德维科或路易吉（Luigi）连在一起使用。他的父亲是都灵公共工程和防务办公室的一名司库，母亲格萝索是一个医生的女儿。拉格朗日是他们生下的七个孩子中的老大，这七个孩子只有两个没有在童年夭折，拉格朗日是其中之一。

尽管拉格朗日的家庭处于意大利社会上层，但却由于投资不当而经常缺钱。家里人决定让拉格朗日学习法律。他考取了都灵学院，在学校里学习法律和古典文学，也开始接触数学，但却发现以欧几里得几何学为主体的数学极其乏味。后来，拉格朗日翻阅了英国天文学家埃德蒙·霍利（Edmond Halley）关于光学的代数方法的书，对数学的看法随之大变。自此，拉格朗日潜心于数学在力学中的运用问题，尤其用功于天体力学，这成了他早期研究中的主要课题。

他娶了自己的堂妹维多利亚·康蒂。“我的妻子是我的一个堂妹，她很长一段时间都跟我的家人生活在一起，是个很好的家庭主妇，从不以虚饰夸人。”他在给同为数学家的好友让·勒·荣德·达朗贝尔的信中如此写道。他曾自陈不想要孩子，并履行了诺言。

拉格朗日在柏林谋得了一个差事，写了大量的研究论文，还获得了几次法兰西学院奖——1772年是与欧拉同获此奖，1774年因对月球动力学的研究获奖，1780年因研究彗星运行轨道的影响效应获奖。他还倾心于数论研究，并证明了这一领域的一个经典问题——四平方和定理。这一定理认为，任一正整数都是四个数的平方的和。例如，$7=2^2+1^2+1^2+1^2$，$8=2^2+2^2+0^2+0^2$，如此等等。

拉格朗日成了法兰西科学院的一员并迁居巴黎，在那里一直待到去世。他认为人们最好遵守自己所在的国家的法律，不管对它认不认同，这一观念让他在法国大革命中免受了其他知识分子罹遭之祸。1788年，拉格朗日发表了巨著《分析力学》(*Analytical Mechanics*)，他把力学当成数学分析的分支来写。他为自己的皇皇巨著中没有任何关于矢量的内容而志得意满，他认为，这可以使逻辑更加严密。

1792年，拉格朗日娶了第二个妻子，天文学家勒·莫尼埃（Le Monnier）的女儿莱妮-弗朗索瓦-阿德莱德。1793年8月，法国正处于恐怖统治时期，法兰西学院被关闭，只保留了计量大会的活动。很多重要的科学家被处决了，其中包括化学家拉瓦锡、物理学家库仑以及拉普拉斯。因此，拉格朗日成了计量大会的主席。

他的意大利国籍给他带来了麻烦。革命政府通过了一项法令，要逮捕敌对国家的外国人。当时还具有相当影响力的拉瓦锡为拉格朗日斡旋，使他得到了豁免。但是没过多久，一个革命小组就宣判了拉瓦锡死刑，在判决的第二天就把他推上了断头台。拉格朗日写道："暴徒在刹那间就能砍掉他的头，但是一百年也不能再生出这样一个人才来！"

在拿破仑统治时期，拉格朗日被授予了好几项荣誉：1808年获得了皇帝荣誉勋章和伯爵头衔，1813年又获得了帝国十字勋章。就在获

得十字勋章的一星期后，拉格朗日去世了。

*

1770 年，也就是发现四平方和定理的那年，拉格朗日开始着手一本关于方程的长篇论文，他写道，“我写这篇论文是要检验迄今为止的方程的代数解法，把它们引入一个普遍原理，对这些方法只适用于三次和四次方程而不能解决更高次方程作出解释。”正如让-皮埃尔·蒂格诺（Jean-Pierre Tignol）在《加洛瓦的代数方程理论》（*Galois's Theory of Algebraic Equations*）一书中所言，拉格朗日的“明确目的是确定这些方法为何如此而不是它们如何发挥作用”。

拉格朗日对这些文艺复兴时期创造的方法的理解比它们的创造者更加深入，而且证明了自己对这些方法为什么不能适用于五次以上方程的设想。但是，他却没能进一步解释是否有能够解出五次以上方程的方法。但是，他告诉我们他的结论“将对那些想要解出更高次方程的人非常有用，给他们提供了支撑这一观点的诸多见解，为他们省去庞杂而无用的尝试”。

拉格朗日发现，卡尔达诺、塔尔塔里亚及其他人所使用的所有特殊技巧是建立在同一思想基础之上的。这一思想并没有直接去求方程的根，而是把这一问题变换为辅助方程的求解问题，这些辅助方程的根必须与原方程的解有关然而又有所区别。

三次方程的辅助方程比较简单，也就是一个二次方程。这个“分解出来的二次方程”用巴比伦人的方法就可以求解；然后，三次方程的解就可以通过开立方求得。这就是卡尔达诺公式的结构。四次方程的辅助方程也比较简单——一个三次方程。这个“分解出来的三次方

程”也可以用卡尔达诺的方法求解，而这个四次方程的解就可以通过一个开四次方求得——也就是两次开平方。这就是费拉里公式的结构。

我们可以想像当时的拉格朗日有多么兴奋。如果这一模式可以推演下去，那么五次方成就会有一个“分解出来的四次方程”，然后就可以用费拉里的方法求出一个五次方根。继而，六次方程就可以分解出五次方程，那么，人们就能用可以名之为“拉格朗日方法”的方法求解了。这样，他就将解出所有的高次方程了。

可是残酷的现实给了他沉重的一击。五次方程的分解方程不是四次方程，而是一个更高次数的方程——六次方程。简化三次和四次方程的方法却把五次方程复杂化了。

数学的进步并不是通过解决越来越难的问题来完成的。拉格朗日验证了这种方法无法适用于五次方程。但是他并没有证明五次方程无法求解；因为很可能存在着另一种行之有效的方法。

难道不是吗?

对拉格朗日来说，这只是个反问句。但是，他的一个继承者却很重视这一问句，并且作出了回答。

*

这个人就是鲁菲尼（Paolo Ruffini），而我说他对拉格朗日的反问“作出了回答”，其实带有些许的欺骗成分。他认为他作出了回答，而且他的同时代人都没有发现他的答案有任何错误——部分是由于他的著作并没有引起人们的足够重视，所以没有得到实际检验。鲁菲尼终其一生都相信自己证明了五次方程无法用根式求解。直到他死后，人们才发现，其实他的证明存在着一个重大的漏洞。在鲁菲尼复杂的推

论过程中，这一漏洞很容易被人们忽略。这一错误是一个“明显”的假设造成的，他却几乎从来没有意识到自己作出了一个假设。

稍微有些经验的数学家都知道，人们很难发现那些不成立的假设，之所以说很难发现，恰恰正因为它是不成立的。

鲁菲尼出生于1765年，是一个医生的儿子。1783年，他进入莫德纳大学学习医药、哲学、文学和数学。他跟着路易吉·凡蒂尼（Luigi Fantini）学习几何学，跟鲍洛·卡西亚尼（Paolo Cassiani）学习推算。在卡西亚尼成为埃斯特家族的继承者，管理巨大的产业之后，身为学生的鲁菲尼就接手了卡西亚尼的数学分析课程。1778年他获得了哲学、医药和外科学位，1789年又获得了数学学位。没过多久，他就接替了视力下降的凡蒂尼的教授职位。

时局变动打断了他的学院生涯。1796年，拿破仑·波拿巴打败了奥地利和萨丁军，又把目光转向了都灵，并且占领了米兰。很快，他又占领了莫德纳，鲁菲尼被迫卷入了政治漩涡。1798年，他想重新回到大学，但是被拒绝了，由于宗教背景的原因，他发誓效忠共和国。由于没有工作，他得以把更多的精力投入到自己的研究中，集中精力研究复杂的五次方程问题。

鲁菲尼认为之所以没有人能解释五次方程无法求解的原因，就是五次方程是不可解的。具体地说，就是根本没有关于比根式更简洁的解五次方程的公式。在1799年出版的两卷本著作《方程的一般原理》（*The General Theory of Equations*）中，他声称可以证明自己的论断，他认为，“四次以上方程的代数解法是不可能存在的。请看下面的重要命题，我可以肯定（如果我没错的话）：给出对这一论断的证明是我发表这本书的主要原因。在不朽的拉格朗日的伟大思考中，已经为这一证明打好了基础。”

这项证明包含了长达500页的蹩脚的数学论证。其他的数学家读这样一本书是个苦差使。除非有充分的理由，否则即使在今天也不会有人读完这样一项冗长而琐屑的证明。如果鲁菲尼提出了一种解法，他的数学界同仁肯定早就试着去证明它了。但要他们把那么多时间花在这样一个消极的结论上，可以想像会有多么强人所难。

尤其要说明的是，这一结果很可能是错误的。在一本500页的数学著作中的第499页发现一个错误，还有比这更让人气恼的事情吗？

1801年，鲁菲尼给拉格朗日送去了一份书稿，很长时间没有收到回音，他又送出了另一份书稿，并加了一个注释："如果我的证明有什么错误，或者某些我认为是新颖的东西根本就不新颖，或者这本书毫无意义，请您为我直接指出来。"但是，仍然音信全无。1802年，他又作出了相同的努力，依旧一无所获。

数年过去，鲁菲尼一直没有得到他认为本该属于自己的认可。相反，暗示他的证明有错的流言却隐隐约约地传到耳中，但没有一个人指出他的错误在哪里，而鲁菲尼也无法为自己申辩。他的证明确实太复杂了，于是他决定着手寻找更简单的证明，这一决定无疑是正确的。1803年，他完成了自己的新证明，他写道："在这部著作中，我希望用严谨而非艰深的推论证明同一个论题。"但是新证明并没有比原来的证明好到哪儿去。世界并不接受鲁菲尼的观点，也不接受他在1808年和1813年发表的进一步证明。他一直试图让数学界认同自己的著作。预测过天王星位置的德朗布尔（Jean Delambre）在1789年对这一数学问题这样总结道："鲁菲尼打算证明五次方程是无法求解的。"鲁菲尼很快回应道："不仅仅是打算，事实上已经证明了。"

说实话，有那么几个数学家喜欢鲁菲尼的证明。其中一位是柯

西，他有一个简单的记录，除了他自己的观点，在他认为有理的地方都给予了肯定。1821 年，他写信给鲁菲尼："你对于方程的一般解法的讨论，是我一直都在关注的课题，在我看来，你证明了四次以上的方程是完全不可解的。"但是这一赞扬来得太晚了。

1800 年前后，鲁菲尼开始在市军事学校讲授应用数学课程。他继续研习医药学，护理社会上的各种病人，从最穷的人到最富的人都有。1814 年，拿破仑失败以后，他成了莫德纳大学的校长，但政治地位依然极其复杂，若非他处世有方、德高望重，拥有可靠的清誉，他的校长生涯必定会非常艰难。

鲁菲尼同时执掌着莫德纳大学应用数学和临床医学的教鞭。1817 年，流行伤寒病，鲁菲尼继续照看自己的病人，直到自己染病。虽然活了下来，但他再也没有恢复健康，并在 1819 年放弃了临床医学的教席。但是，他从没有放弃科学研究，并于 1820 年发表了一篇基于他同时身为医师和病人的经历的关于伤寒的科学文章。他死于 1822 年，仅在得到柯西对他关于五次方程的工作的赞扬的一年之后。

*

鲁菲尼的著作很难被人接受的另一个原因可能是它的创新性。与拉格朗日相似，他的研究建基于"排列"（permutation）的概念之上。排列是一种将有序串列进行重新组织的方式，大家最熟悉的例子就是洗一副纸牌。它的一般目的是获得一个随机的——无法预测的——序列。一副纸牌的排列方式数量巨大，因此，预测这个随机的洗牌结果几乎是不可能的。

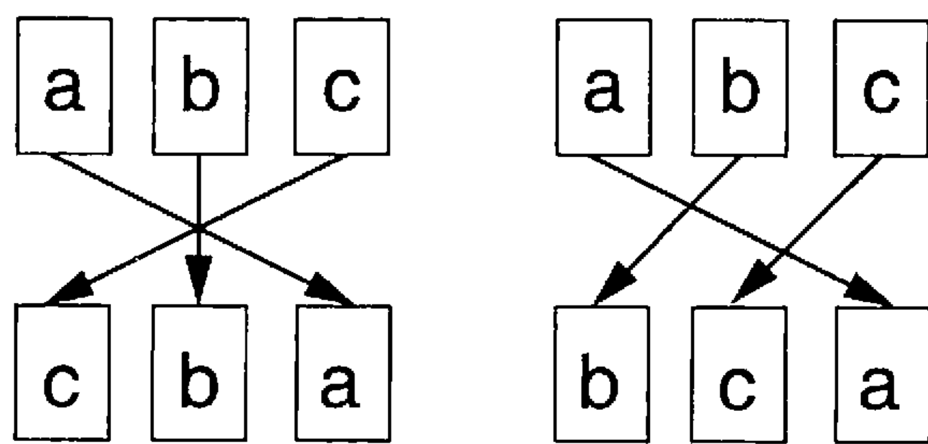
在方程中提出排列理论是因为一个给定的多项式的根可以看做一

个串列。方程的一些非常基本的性质与这些串列的“洗牌”效果直接相关。直观地看，方程“不知道”这些根的排列方式，怎样排列这些根都不会有多大的差异。然而实际上，方程的参数应该是用根表示的完全对称的表达式——不因重新排列这些根而发生改变的表达式。

但是，正像拉格朗日所理解的，一些根式表达式正是一些对称的排列组合。这些“部分对称”的表达式与所有的解方程公式密切相关。鲁菲尼等人对排列的这一性质非常熟悉。但是鲁菲尼却无法系统而熟练地运用拉格朗日的另一种思想：可以轮流将这些排列两两“相乘”，从而得到另外的排列。

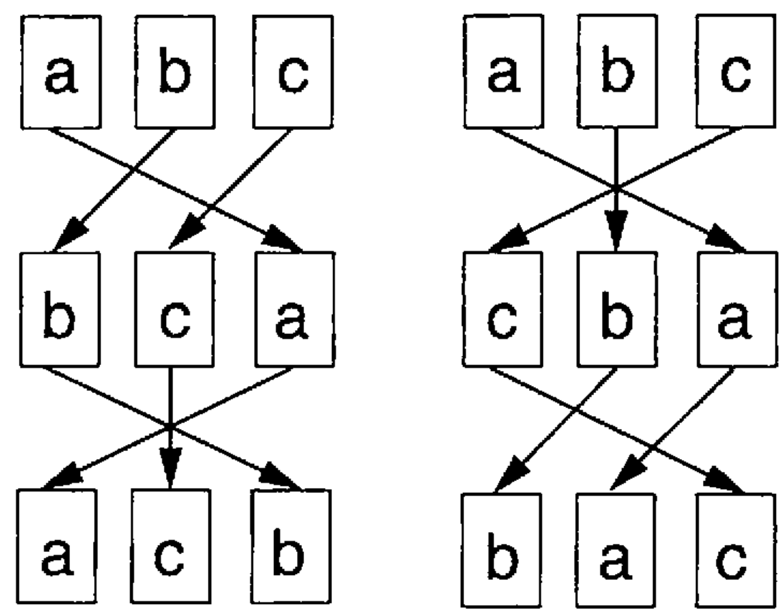
比如 *a*，*b*，*c* 这三个符号，可以得到六种排列：*abc*，*acb*，*bac*，*bca*，*cab*，*cba*。拿出一个排列，比如 *cba*，一看便知，这是一个由这三个符号组成的有序串列。但是，我们也可以将其看成应用于原始串列 *abc* 的一个规则。这样，这一规则就是“颠倒顺序”。这一标准不仅可以应用于 *abc* 这一串列，而且可以应用于所有串列。将之应用于 *bca*，就会得到 *acb*。这样 $cba \times bca = acb$ 就讲得通了。

这一思想在我们这本书中非常关键，如果用图表的形式表示，也许更容易理解一些。我们把 *abc* 重新排列为 *cba* 和 *bca*，得到下面两个图表。



符号 *a*，*b*，*c* 的两组排列

我们可以将这两个组合上下累叠，合并成一个，就得到下面两种形式的组合：



将排列相乘，所得结果取决于第一个串列

我们可以看出，两个排列相乘的结果就是最下面的那个串列，在左图中就是 *abc*。通过这种特定的相乘（不同于一般观念上的数与数的相乘），我们就可以理解 $cba \times bca = acb$ 这一命题。条件就是把第一个排列当做这一堆栈的基础。如果把堆栈中的两层换成不同的串列，就会得到不同的结果。如右图所示，是两个顺序相反的串列相乘，得到的结果是 $bca \times cba = bac$。

*

鲁菲尼关于五次方程不可求解的证明的本质，是提出五次方程可以求解的条件，这个条件就是，五次方程的根可以用根式表达。如果一般的五次方程不能满足这个条件，就没有这一类型的解，因此也就不能沿用解三次、四次方程的解法来求解。

鲁菲尼比着拉格朗日的葫芦画瓢，他熟谙方程根的对称性及它们的排列关系。五次方程有五个根，这五个根一共有 120 种排列。鲁菲

尼意识到，这一系列的排列会获得一定的结构性，而这种结构性，就来自一种有可能存在的解方程公式。如果这种性质不存在，公式也就不会存在。这就像在泥泞的丛林里猎虎，如果真的有老虎，那它就一定会在泥地上留下虎迹；没有虎迹，就没有老虎。

通过研究这种新型乘法的数学规律，鲁菲尼证明了——起码他自己是满意的——这 120 种排列的乘法结构并不具备使方程能用根式求解而必然存在的对称性。他也确实取得了一些实质性的进展。在鲁菲尼开始研究五次方程以前，世界上每一位数学家几乎都认为五次方程是可解的，只剩下如何求解的问题。高斯是个例外，他暗示五次方程是无解的，但是他也认为这是一个无趣的问题，这次他的直觉失败了，而对于高斯来讲，这是很少有的现象。

鲁菲尼之后，人们似乎普遍感觉到，五次方程不能用根式求解。虽然很少人认为鲁菲尼成功地作出了证明，但是人们还是对根式的有效性产生了很大怀疑。这种观念的转变带来了负面影响：数学家们开始对这一问题感到索然无味。

极具讽刺性的是，鲁菲尼的著作有一个重大的漏洞，当时却没有一个人把它指出来。从某种意义上说，他的同时代人的怀疑精神后来有所缓和。但真正的突破只是方法上的：鲁菲尼运用了正确的战略，却没有找到恰当的战术。要解决这一问题，需要一位对战术细节也一丝不苟的谋略家。于是，下面这位人物出场了。

*

1784 年，任劳任怨地在挪威一些最贫穷的偏远山区做了几年出色的牧师工作之后，汉斯 · 马梯阿斯 · 阿贝尔（Hans Mahias Abel）得到

了应得的报偿。他被派往挪威南部海岸的耶尔斯塔教区，距奥斯陆峡湾不远。虽然耶尔斯塔并不十分富庶，但比起他之前管理过的地方强多了，他的家庭经济状况得到了巨大改善。

和以前一样，阿贝尔牧师的任务是精神性的，他要努力让他的教民保持快乐和善良。他来自一个小康之家，曾祖父是一个丹麦商人，做着为挪威军队供货的生意，很赚钱。他的父亲也是一个商人，当过卑尔根（Bergen）的高级市政官。汉斯自尊而谦虚，并非聪明过人，但也绝不愚蠢，精于讨价还价。

为了在贫穷的教区果腹，他在自己的农场上引种了新式植物：纺织用的亚麻和一种名为“地苹果”的根茎蔬菜，也就是马铃薯。他还写诗，与妻子伊丽莎白融洽地一起生活。他家的饮食很出名，从来都没有酒。饮酒是挪威社会的一个大问题，牧师必须为他们的教民树立榜样。但是他有一次还是喝得醉醺醺的来到教堂，向信众展示醉酒是件多么丢人的事情。他只有两个孩子：女儿玛加丽塔和儿子索伦，在当时，这是很少见的小人口家庭。

玛加丽塔资质平平，一生未婚，她的大部分时间都是和父母一起度过的。索伦则完全不同：他反应敏捷、颖悟过人，而且很有创造力，也很有上流社会的品味。他没有父亲那份持重沉着和尽职尽责，并且深受其苦。但他还是子承父业，先是当上助理牧师，后来当了牧师，与一个家族世交的朋友的女儿安妮·玛丽·西蒙森结了婚，并在西南沿海的芬岛（Finnøy）供职。“这里的人很迷信，却满脑子圣经的知识，”他写道，“他们歪曲上帝的权威，从而支持所有荒唐的观点。”尽管如此，他还是很喜欢这份工作。

1801 年，索伦给朋友写信说：“家里又添了喜事一桩，我的妻子在圣诞节的第三天给我生了个胖小子。”这个孩子就是汉斯·马梯阿

斯。1802年夏天，他的弟弟尼尔斯·昂利克出生。从生下来那天起，尼尔斯就一直生病，母亲不得不花大量的时间来照顾他。

欧洲的军事情势日趋紧张，联合了的挪威和丹麦被夹在强大的法国和英国之间。出于自己的考虑，拿破仑想与之结盟，所以，在英国与瑞典达成协定后，挪威—丹麦立刻变成了英国的敌人，遭到了英军的入侵。三天后，为了保护哥本哈根不受破坏，挪威—丹麦联盟投向了拿破仑。后来，拿破仑的军事控制开始减弱，他的助手贝尔纳多特变成了瑞典的国王。在挪威向瑞典投降后，挪威议会被迫接受贝尔纳多特为国王。

*

1815年，马梯阿斯和尼尔斯兄弟俩被送到奥斯陆的教会学校。数学教师彼得·巴德尔是那种通过身体暴力督促学生的人。但是，这兄弟俩的成绩很好。1818年，巴德尔竟然一下把自己的一个学生、一个国会议员的儿子打死了。令人吃惊的是，巴德尔竟然未被判刑，但是他的数学教职还是被当过数学教授克里斯托弗·汉斯丁（Christoffer Hansteen）助手的伯恩特·洪波义（Bernt Holmboe）接替了。这成了尼尔斯数学事业的转折点，因为洪波义允许学生们钻研大纲之外的有趣问题。尼尔斯获准借阅经典教科书，其中就有欧拉的著作。“从那时起，”洪波义写道，“（尼尔斯）阿贝尔怀着强烈的求知欲投入到数学当中，并以天才特有的速度进步着。”

在离开这所学校前不久，尼尔斯确信自己已经解出了五次方程。洪波义和汉斯丁都没有发现任何错误，因此，他们把尼尔斯的运算过程交给了丹麦著名数学家费迪南德·德根（Ferdinand Degen），希望能够通过丹麦科学院发表。迪甘也没能在运算过程中发现任何错误，但

作为对解五次方程略知一二的老手，他要求尼尔斯进行举例运算。尼尔斯很快意识到某些地方出了差错，但是又庆幸自己没有因发表这篇文章而闹出笑话。

索伦野心勃勃，但不够世故，这让他陷入了窘境。他说两个挪威议会代表不公正地关押了一名铁艺经理人，而这些铁艺就是这两人中的一位的财产。这一指控引起了轩然大波。但是后来真相大白，事实证明索伦的话是不可信的，但是索伦拒绝道歉。他灰心丧气、郁郁寡欢，终于酗酒致死。在葬礼上，索伦酩酊大醉的遗孀安妮·玛丽把自己钟爱的仆人拉上了床。第二天早上，几名公差来访，她却还在床上，而他的情人就躺在她身边。她的一位姑母写道："这可怜的孩子，真让人伤心。"

1821年，尼尔斯从教会学校毕业，并被克里斯蒂安尼亚大学（University of Chritiania，如今的奥斯陆大学）录取。他的算术和几何都获得了甲等，其他数学学科的成绩也非常好，而且每门功课都极其出色。由于经济窘迫，他申请了免费宿舍和生火的木柴，还得到了一项维持日常生活的资助。一些教授发现了他过人的天赋，出钱让他做助手。有了这些资助，阿贝尔一心投入到了数学研究中，开始钻研五次方程，完善此前自己流产的努力。

*

1823年，尼尔斯开始研究椭圆积分，对这一领域的研究一直持续到他生命的最后时刻，成就甚至高于他在五次方程方面的研究成果。他试图证明费马最后定理，尽管举例说明了要证伪费马最后定理要用到超级大数，但他既没能证明也没法证伪。

这一年夏天，他参加了一次舞会，遇到一位年轻女子，并邀其跳

舞。但是，好几次都没有跳好，最后，两个人大笑起来，因为他们两个人都不怎么会跳。这名女子就是克里斯汀·坎普（Christine Kemp），人们通常叫她“克里丽”（Crelly），一名战争委员会委员的女儿。和尼尔斯一样，她也没钱，靠做家教为生，从日常琐事到科学什么都教。“她不漂亮，红头发，有雀斑，但他是个很好的女孩，”他写道。他们相爱了。

这件事促进了尼尔斯的数学事业。1823年快要结束时，他证明了五次方程不可解——和鲁菲尼的临门失脚不同，他的证明完美无瑕。尼尔斯的证明框架和鲁菲尼相似，但是技巧上更高一筹。起初，尼尔斯并不知道鲁菲尼的著作。他后来一定了解了鲁菲尼的著作，因为他曾间接提到过它的缺陷。尽管尼尔斯也没有讨论过鲁菲尼著作的缺陷具体在哪儿，但事实上他的工作正是弥补鲁菲尼漏洞的桥梁。

尼尔斯和克里丽订婚了。他们订婚后，尼尔斯不得不去工作，这意味着他必须要让欧洲的大数学家都了解自己的天赋。仅仅发表他的论文是不够的，他得在这些太岁们头上动动土，但这需要一大笔旅行费用。

几经周折，克里斯蒂安娜大学同意给尼尔斯一笔费用，让他到巴黎访学，在那儿他将见到世界顶尖的数学家。在准备出行期间，他决定将自己最好的研究成果做几份拷贝。他相信他对五次方程不可解的证明能够引起法国数学界的重视。很可惜，他的作品都是在一家小报用挪威文印刷的。因此，他决定私自把自己关于方程的研究成果印成法文。他为之取名为“论代数方程，五次方程无法用一般方程的解法求解”。

为了节省印刷费用，尼尔斯只好把自己思想的核心浓缩出来，竟然只有六页纸的长度。这可比鲁菲尼的500页少多了，但是在数学中，有时简短会使论述变得晦涩难懂。一个论题中的一些十分关键的逻辑

过程被尼尔斯省略了。这本拷贝成了一个提纲，而不是一个证明。

他写了一段话介绍自己的证明："数学家们总是忙着寻找解代数方程的普适方法，也有人想要证明它不可求解。我希望，这篇文章会对他们有所助益，因为它填补了方程理论的一项空白。"然而，这是一个近乎无望的希望。虽然他在巴黎成功拜访了几位数学家，也让他们看了自己的论文，但是其中的论证过于简化，大多数数学家都没能看懂。高斯也收到了一份，但压根就没有读过，人们在死后发现了它，连封都没有拆。

后来，也许是尼尔斯意识到了自己的错误，所以又做了两份较长的证明，添加了更多的细节。而此时他听说了鲁菲尼，因此在这一版本里写道："第一个尝试证明（五次方程）无法用一般方法求解的是鲁菲尼，但是他的论文太过繁复，很难确定其论点的正确性。在我看来，他的很多论证难以让人满意。"但是和其他人一样，他也没有说出为什么。

*

鲁菲尼和阿贝尔在论证过程中使用的是当时程式化的数学语言，不太适合用来思考此类问题。数学家们大多只看到了一些细节和具体观念，但是这两人的方程理论是应该用更具一般性的术语进行思考的——它更多的是关于结构和过程而不是细节。因此，他们超越时代语言的思想很难被同代人理解。即使在当今，用当时那套术语去理解他们的思想也是很难的。

幸运的是，我们可以通过引入建筑的比喻来理解他们的分析的基本特征。理解鲁菲尼的"差不多证明"和阿贝尔的完全证明的方法之

一，就是将其想像成一座塔的修建。

这座塔每层只有一个房间，通过一道梯子与上层的房间相连。每个房间里有一个大口袋。打开口袋，就会蹦出成千上万的数学公式，铺满整个房间。乍一看，这些公式没有什么特殊结构，就好像是很多从数学教科书中拿来的公式杂乱地堆在一起，有的长，有的短，有的简单，有的又极其复杂。但是靠近看时，就会发现其家族相似性(family resemblance)，一个袋子里的公式都有很多共同特征。我们爬得越高，袋子里的公式就越复杂。

第一层袋子里的方程都是基础性的，可以带入方程参数做加减乘除，就这样反复地做，想做多少次就做多少次。在代数方程领域，只要有了参数，所有此类“无害的”合成（运算）都可以轻易完成。

顺着梯子往上爬，你就必须从袋子里拿出几个方程，用它们建立一个根式。它会是个平方根、立方根，或者五次方根，或是其他此类的东西。但是你所要求其根的公式必须是从袋子里拿出来的。你可以把它看成一个 P 次方根，将 P 作为一个基点，因为更复杂的根可以从这一基点上建立起来，这将对你很有帮助。

无论你想求一个几次方根，在你来到第二层时都会发现另一个袋子，里面装的东西和第一个袋子里的东西最开始是一致的。打开袋子，把你的新根式扔进去。

公式繁殖了。当诺亚的方舟停在亚拉腊山上，他告诉方舟上的生物出去并开始生育。这些公式可不仅仅繁殖，他们还会加减乘除。在几秒钟的疯狂活动之后，第二层的袋子里鼓鼓囊囊地装满了方程参数和你的新根式的所有可能的、“无害的”合成物。与第一个袋子相比，第二个袋子里有了很多的新公式，每个新公式都是相似的，因为它们都以你的根式为新的组成部分。

在第二层重复同样的做法就可以到达第三层。从新口袋里（一层只有一个口袋）拿出几条公式，通过求根得出一个根式。把你的新根式带到第三层，扔到口袋里，就让公式们杂交去吧。

就这样，在每一层引入一个根式，口袋里就会产生一个新公式。无论哪一步，所有的公式都是以参数和所引入根式为基础建立起来的。

最后你会到达塔的最高层。你的求索（用根式解方程）完成了，把顶楼的口袋翻个底朝天，你至少能找到这个方程的一个解。

可能存在着很多座合理的塔，它们会随着你选用的公式和根式的不同而改变；当然也有很多一败涂地的情况，连方程根的影子也摸不着。但是，只要问题有解，只要以根式建立的公式能得出一个解法，在相应的塔的顶层就会有方程的一个根，因为，这些公式会准确地告诉我们怎样通过一个又一个的根式求得方程的根，也就是说，这些公式准确地告诉了我们该怎样建塔。

*

我们可以用建塔的理论重新解释三次、四次方程的经典解法，甚至是巴比伦的二次方程解法。我们之所以以三次方程开始讲，是因为它足够典型，又容易理解。

卡尔达诺的塔只有三层。

第一个袋子里装着很多参数及其混合物。

要登上第二层需要一个二次方根。这个二次方根是特别的，只能是通过第一层的袋子里的根式产生的。第二层的袋子里装满了与这个二次方根有关的混合物以及参数。

现在去第三层，也就是顶层，需要一个三次方根，同样也是很特

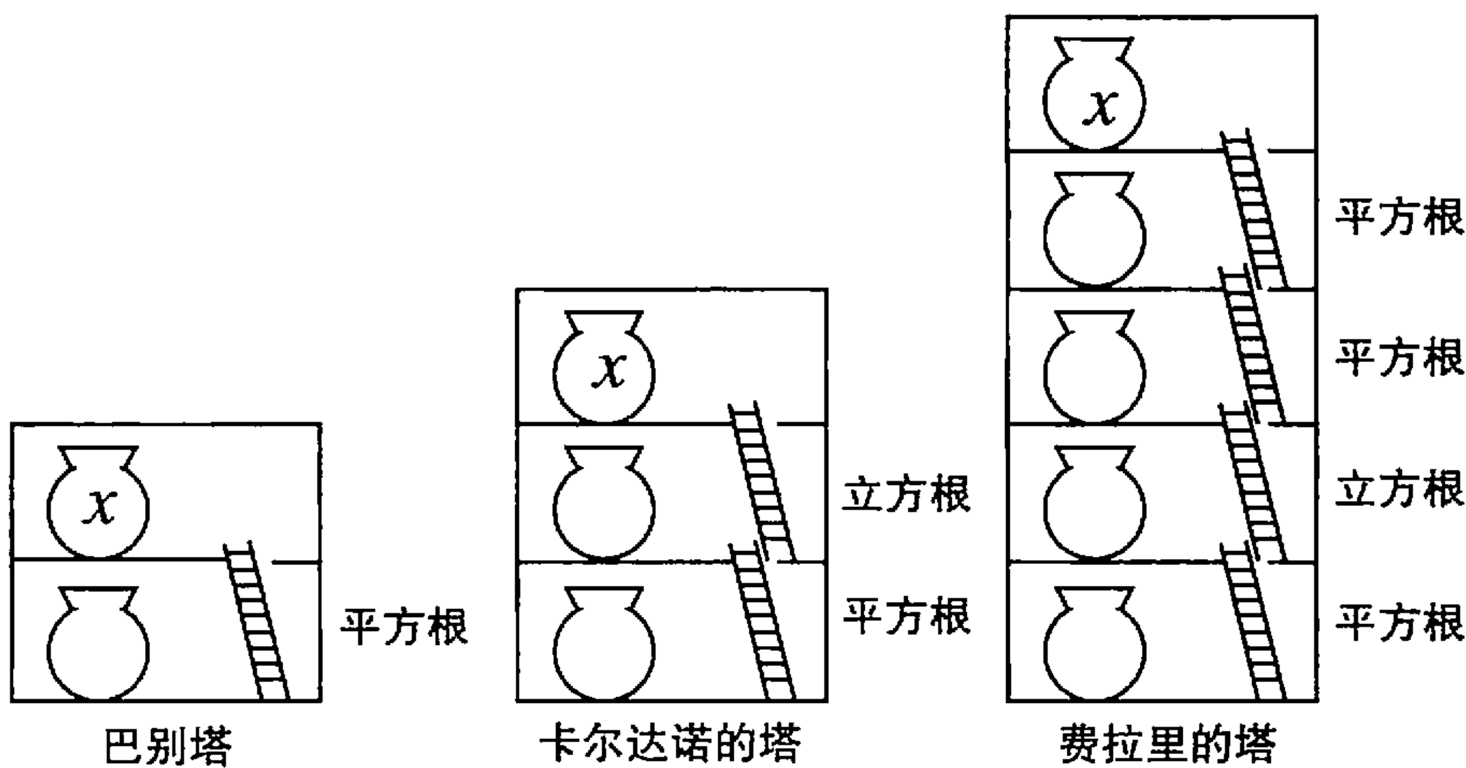
殊的一个。这个三次方根是由参数和你曾借以登上第二层的二次方根产生的。第三层的袋子里到底有没有三次方程的解呢？当然有，卡尔达诺的公式就可以证明。登上塔顶的过程就是一步步取得胜利的过程。

费拉里的塔更高些，有五层。

第一层的袋子里通常只有参数的合成物；通过制造“无害的”合成物进入第二层，然后得出一个合适的二次方根；再通过制造无害的合成物来到第三层，再得出一个立方根；再通过制造无害的合成物到达第四层，再得出一个合适的平方根；最后再通过制造无害的合成爬上最上面的第五层，然后得出一个合适的平方根。

现在，你所要的五次方程的根已经在顶层的袋子里了。费拉里的公式正是如何建这样一座塔的恰当说明。

解决了二次方程的巴别塔也适用于这个比喻，但只是一座只有两层的矮塔。第一层的袋子里只有参数的混合物，从中仔细挑选一个根来到第二层，也就到了顶层。这一层的袋子里是一个二次方程的根，事实上，这两个根都是二次方程的根。你在学校学到的解二次方程的公式的创造者巴比伦人就是这样告诉我们的。



解二次、三次和四次方程

*

五次方程到底是怎么回事呢?

假设用根式解五次方程的公式确实存在，而我们不知道它到底是什么，但我们可以作出很多与之相关的推测。尤其是，它肯定对应于某座塔。我们把这座假想之塔叫做巴别塔好了。

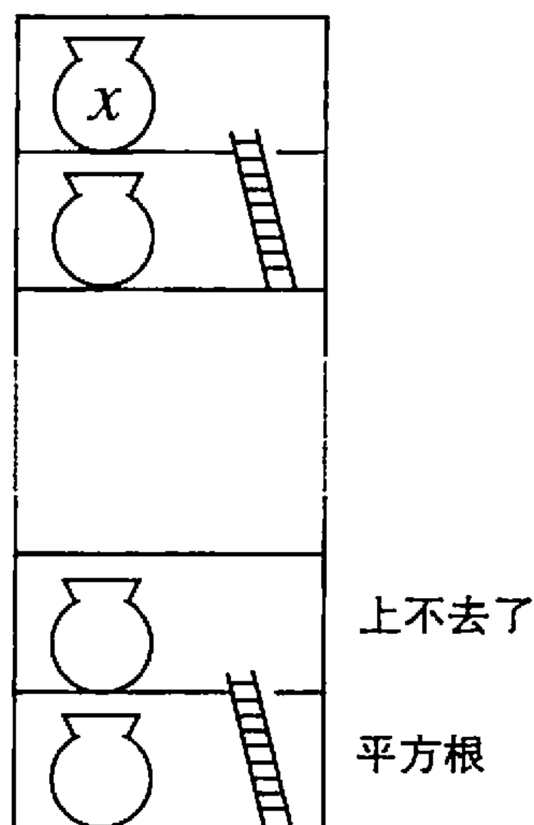
巴别塔可能有成百上千层，而它的梯子跟各种各样的根式相关——19 次方根、37 次方根，我们对此并不很清楚。我们可以肯定的是，第一层的袋子里只有参数的无害混合物。我们可以遐想一下，在它高入云端的顶楼的袋子里有一个五次方程的解。

我们想知道怎样攀登这座塔，数学告诉我们只有一条路通往第二层。我们只能选出一个平方根，别无他途。

事实却不尽然如此。我们可以讨论其他各种各样的根，建一座高大的塔。但是如果不是在塔的某一层有一个对应的平方根，塔顶就不会有你要求的根。而且，从之前的任何一层都不能到达塔顶，建这样一座塔实在是浪费时间和金钱。明智的建造者在一开始就要全力以赴地寻求这个平方根。

爬上第三层时你需要的是什么呢?

没有往第三层去的梯子。你可以登上第二层，然后就只能止步不前了。如果你到不了第三层，也就当然无法找到袋子里的答案了。



阿贝尔的塔

为什么五次方程不可解

简言之，巴别塔是不存在的，

存在的只有终结于第二层的种种尝试，也可能是很多精巧而无用的多层建筑物，但也注定要以同样的方式终结，其失败的原因都是一样的。这就是鲁菲尼证明了的，虽然他在技术上存在着缺陷。大致可以说，他其实没能证明，因为根式的混合物有可能存在于顶层，根式本身也一样。

鲁菲尼的证明和阿贝尔的塔有着明显的相似性。但是通过建塔的形式，阿贝尔发展了鲁菲尼的技巧，并弥补了他留下的缺陷。他们证明了没有从方程参数到方程根的根式之塔。通过建筑学的语言，这告诉我们，没有只用根式就能解五次方程的公式。用根式解五次方程就像想要站在自己肩膀上爬上月亮一样，毫无可能。

*

1828年的圣诞节就要到了，阿贝尔准备与自己的朋友凯瑟琳和尼尔斯·特里肖一起在芬罗兰（Froland）度过。他想去探访克里丽，她就住在附近。鉴于阿贝尔的健康状况，他的医生不赞成这次旅行。在给克里斯托弗·汉斯丁的妻子琼尼的一封信中，凯瑟琳写道："如果你当时还在镇上，他就可能会愿意留下。可是他极力掩饰着自己的病情。"12月中旬，阿贝尔动身前往芬罗兰，为了御寒，他把自己裹得严严实实。12月19日到达时，他几乎把所有的布片都裹在了身上，甚至把袜子都穿到了手臂上。尽管他不断咳嗽，冷得发抖，但还是在数学的园地上不懈耕耘，在特里肖的起居室被他们的孩子围绕着快乐地工作着。他很喜欢这些小伙伴。

阿贝尔还在努力谋求一个体面的职位，可是就连他在奥斯陆的临时职位也已不保险了。圣诞节过后，他把主要精力用于保住自己在

柏林的工作上。他的朋友奥古斯特·克莱尔也在背后为他奔走，他说服教育部成立了一个数学研究所，希望能在那里为阿贝尔谋得个教授职位。他得到了科学界的大人物亚历山大·冯·洪堡的支持，另外还有高斯和法兰西学院的阿德利昂-玛利·埃·勒让德（Adrien-Marie Legendre，1752—1833）的推荐信。克莱尔告诉教育部长说，阿贝尔正准备接受柏林的一个职位，所以当局必须赶快行动，因为还有很多地方想要阿贝尔去工作，尤其是哥本哈根。

阿贝尔原打算在 1 月 9 日离开芬罗兰前往奥斯陆，但是他的咳嗽和发抖都更严重了，大部分时间只能待在自己的屋子里。原来是他未来岳父母的坎普两口子非常担心他。在他原打算动身的那天早上，他剧烈地咳嗽，而且咳出血来。家庭医生马上被叫来了，他嘱咐阿贝尔卧床休息，自己则一直照料着他。克里丽像个护士一样，她温情悉心的照料让阿贝尔的病情了有了明显的好转。几个星期后，阿贝尔获准可以每天在椅子上小坐一会儿，但不准进行数学研究。

勒让德写信给阿贝尔，告诉他自己对他的椭圆公式印象深刻，并建议这位年轻人发表解决如何确定方程是否能用根式求解的方法："我建议你尽快将这一理论付梓，这会给你带来极大的声誉，大家会将之看成关于这一长期存在的数学问题的最伟大发现。"很多著名数学家，有的直接阻挠阿贝尔影响深远的著作的发表，有的又置之不理，而他名声却在另外的圈子中迅速崛起。

1829 年 2 月底的时候，阿贝尔的医生发现他再也无法康复了，只能希望尽可能久地保证病情不再恶化。医生给阿贝尔从前的老师洪波义发了一份医疗证明，报告这位年轻人的健康状况：

> ……他来到芬罗兰不久就患上了感冒，很快又引发了肺炎咳

血。由于连续性的咳嗽和极度虚弱，直到现在他还只能躺在床上休息，另外，他连很微弱的气温变化都承受不了。

更严重的是，干咳引起的胸腔震痛很可能导致隐蔽性肺病和支气管结核，这些又会引发随之而来的胸痨，特别是他现在所处的环境，更容易引发这类病症。

由于病情很危险……恐怕在春天到来以前他回不了奥斯陆了。到那时，就算无法负担自己的医疗费用，他也必须回去辞掉自己的职务。

克莱尔是在柏林得知这一坏消息的，为保住阿贝尔的职务，他拼尽了全力，并建议德国有关部门把阿贝尔转移到环境更好的地方。

4 月 8 日，从克莱尔那里传来了阿贝尔受到提携的好消息：

教育部决定让你去柏林赴职……你会获得一个与你的能力相称的工作，但我无法告诉你薪水是多少，因为我也不知道……我希望让你尽快得知这一消息，你应该被照顾得不错吧。为自己的未来着想，不要再有什么顾虑了。你肯定会成为我们中的一员。

然而事与愿违。

阿贝尔已经病得无法走动了，只得待在芬罗兰，尽管有克里丽的精心照料，他还是越来越虚弱，咳嗽得越来越厉害。只有在更换席子的时候，他才能够下床。当他想要研究数学问题时，却发现无法执笔。他贫病交加，只能沉湎于过去之中，但他从不向自己的爱人吐露自己的感受，一直到大限将至都保持着良好的脾性。

作为阿贝尔的未婚妻，克里丽很自然地发现已经难以掩饰自己的

痛苦，玛利或汉娜就让她守在床前。阿贝尔咳嗽得越来越厉害，已经无法入睡。汉娜一家雇了一个护士照看阿贝尔，这样，克里丽可以休息一会儿。

4月6日，在被发烧折磨了一夜之后，阿贝尔去世了。汉娜写道："4月5日晚间，他经受了最大的痛苦。天快亮时，他越来越安静了，上午11点，他撒手人寰。我的妹妹和他的未婚妻伴他度过了最后时刻，看着他安静地投入了死神的怀抱。"

五天后，克里丽写信给凯瑟琳·汉斯丁的妹妹亨利耶蒂·菲德里克森，让她通知凯瑟琳发布这一噩耗。"是的，我最亲爱的人，我不得不写下这封信，我很感激您的姐姐。我用颤抖的笔请求您告诉她，她失去了一个善良、虔诚、对她无限热爱的年轻人。

"我的阿贝尔死了！……我失去了世上的一切，我已经一无所有。原谅我，我不能再哭诉不幸了。告诉她收存我包好的阿贝尔的头发留念吧，要以最仁慈的方式把这件事告诉你的姐姐。你不幸的克里丽·坎普。"

倒霉的革命家

数学是永不止步的。一个问题的解决就是另一新问题的诞生。阿贝尔死后不久，他关于有些五次方程无法用根式求解的证明开始被人们接受。但是阿贝尔的著作只是个开始。尽管之前所有解全部五次方程的尝试已经渐次停下了，但是一些聪明的数学家还是证明了一部分五次方程可以通过根式求解。当然不止最明显的 $x^5-2=0$，$x=\sqrt[5]{2}$，还有一些比较令人吃惊的五次方程，比如 $x^5+15x+12=0$，但是求解方法太复杂，在此就不作详述了。

这是个难题。如果一些五次方程可以用根式求解而另一些不能，那么，它们之间的区别到底是什么呢？

这一问题的答案改变了数学和数学物理学。虽然这一答案已经出现了 170 年之久，但直到现在还在不断产生重要的新发现。回过头来看，这真是令人惊叹。解五次方程看起来没有什么实际用途，如果在工程学或天文学中遇到了五次方程，会有很多办法帮助我们找到解法，就像在很多我们迫切需要的领域中一样。五次方程可不可以通过

根式求解是个典型的“纯”数学问题。

这样想你就错了。

阿贝尔已经发现了某些方程用根式求解的障碍。这些障碍使这些方程根本不存在用根式求解的可能。而那些爱打破沙锅问到底的数学家总会纠缠于那些主要问题解决之后仍不得解决的问题：“不错，我明白——但是，原因是什么呢？”而这，也正是我们本章的故事所要阐明的问题。

这种态度看起来没什么好处，但事实一次又一次地证明，这非常重要。之所以说这种态度没有好处，是因为很多数学问题太难了，没人能够解决。所以，当有人试图解决前人未能解决的问题时，仅仅找到一种解决方法是远远不够的。找到这种方法可能是因为幸运（数学家是不屑于这种幸运的）或者某些特殊原因。如果这一原因能够解释，很多其他问题就能用相似的方法解决。

所以，当阿贝尔钻研“是不是所有的五次方程都是可解的？”这一问题并得到完全否定的答案时，另一个思考者已经开始思考另一个更具普遍意义的问题了：哪些方程可以用根式求解，哪些不能？平心而论，阿贝尔已经沿着这一方向开始思考了，如果不是结核夺去了他的生命，他可能已经找到了答案。

*

这个改变了数学进程的人就是艾瓦里斯特·加洛瓦。他是数学史上最具传奇色彩和悲剧色彩的数学家之一，而他的一大批发现却差一点遗失一空。

毫无疑问，即使加洛瓦没有出生，其他人应该也已经得出了同样

的发现。很多数学家已经开始在这一领域展开了探索，只是与那些重大发现失之交臂。也许在另一个世界，某个具有加洛瓦的天赋和见解的人（或者阿贝尔能在结核病中多活几年）也可以完成同样的思想历程。但是在现实世界中，这个人只能是加洛瓦。

加洛瓦1811年10月25日生于巴黎远郊的堡拉瑞恩（Bourg-La-Reine），当时，堡拉瑞恩还只是一个小村庄。现如今，它已经成了隶属于塞纳区的近郊，就在N20高速公路和N60高速公路的交汇点上。N60高速公路现在被称为加洛瓦大道。1792年，堡拉瑞恩村曾被命名为布尔－厄加利特（Bourg-l'Egalite），这反映了当时的政局动荡："皇后镇"已经变成了"平等镇"。1812年，这个村庄的名字又变成了堡拉瑞恩，而大革命的气息仍然挥之不去。

艾瓦里斯特的父亲尼古拉斯－加布里埃尔·加洛瓦是一个共和主义者，也是村庄自由党（厄加利特解放阵线）的领导人，主要政治主张是废除君主政体。1814年，通过连蒙带骗的妥协，路易十八重掌大权，尼古拉斯－加布里埃尔变成了镇长，从政治信念上说，在这样一个政府中，他感到很不自在。

艾瓦里斯特的母亲阿德莱德－玛丽生于德芒特（Démante）家族。她的父亲是一位法学学者，一个专业律师帮办，职责是就法条问题发表意见。阿德莱德－玛丽精通拉丁语，并把自己的古典学知识传授给儿子。

艾瓦里斯特生命中最初的12年是在家里度过的，接受母亲的教育。本来，10岁时兰斯的一所学院准备让他入学，但是母亲觉得10岁离家太早了点。12岁时，他进入路易大帝高中（Collége de Louis-le-Grand），一所预科学校。艾瓦里斯特到学校不久，学生们就开始拒绝在学校的礼拜堂唱颂歌，小加洛瓦亲眼看到了革命的前兆，学校立

即开除了100名学生。多亏了数学课，他逃过了这次惩罚。

入学后的头两年，他获得了第一个拉丁文的奖励，但他很快就厌倦了。后来，为了成绩，学校强迫他重复自己学过的课程，这进一步加深了他的厌倦，情况变得越来越糟。数学把加洛瓦从下坡路上拉了回来，这门学科的丰富内容足以笼络加洛瓦的兴趣。加洛瓦没有把时间花在其他数学问题上，而是直奔经典著作——拉格朗日所编的《几何原本》，这有点像现在的物理学学生从一开始就去读爱因斯坦的著作。但是数学中有一种门槛效应，有一个智力极点。如果一个学生战胜了最初的几道难题，理解了这门学科的符号特性，掌握了理解数学思想的最佳方式，而不是机械地学习，他就会轻松地驶入快车道，奔向更复杂更抽象的问题，只有那些愚钝的学生才会被等腰三角形之类的几何问题难住。

加洛瓦能不能理解拉格朗日言简而意繁的著作一直是人们争论的问题，但不管怎么说，他没有被吓倒。他开始阅读拉格朗日和阿贝尔的专业论文，他们研究的领域很自然地变成了加洛瓦兴趣的中心，尤其是方程理论，方程问题可能是惟一真正吸引了加洛瓦的东西。数学大大占用了加洛瓦日常功课的时间。

在学校，加洛瓦自由散漫，这一习惯他一直都没有改掉。他为难自己的老师，只在头脑里解决问题，而不是通过作业写出来。在今天，强迫写作业是老师折磨有天分的学生的法宝。我们可以试想一下，一个崭露头角的足球运动员每踢进一个球，教练就要求他按顺序写下自己踢进这个球所使用的技巧有多么痛苦。而事实上，根本就没有这样一串乱七八糟的东西，每个了解游戏规则的人都知道，有空当就要把球踢进去。

我们年轻的数学家完全能够做到。

加洛瓦野心勃勃：他想到法国最有名的巴黎高科（École Polytechnique）继续学习，那里是法国数学家的孵化器。但是他没有接受数学老师的建议，拒绝系统的学习方法，拒绝交作业，拒绝针对考试进行学习。最后，艾瓦里斯特参加了入学考试，可是由于准备不足和过分自信，他没有考上。

20年后，一位颇有影响的法国数学家欧利·特尔奎姆（Orly Terquem），在一份著名的期刊上解释了加洛瓦失败的原因："一位高级知识的候选人会由于次级知识的问题被淘汰。因为他们不了解我，我是个不开化的人。"现代的评审者更注重沟通技巧，并不那么要求一个高级知识的候选人必须通过次级知识的测试。加洛瓦毫不妥协的个性对自己没有任何帮助。

因此，加洛瓦只能继续待在路易大帝高中，在那儿，他偶尔还会获得人们的赏识。一个叫路易-保尔·理查德的老师发现了加洛瓦的天赋，而加洛瓦也开始跟着他学习高等数学的课程。保尔·理查德认为，以加洛瓦的天分，他应该免试进入巴黎高科，但是理查德也认为加洛瓦不会通过巴黎高科的入学考试。我们现在没有理查德曾向巴黎高科阐明自己观点的证据，即使他这样做了，也没有人会注意。

*

1829年，加洛瓦发表了自己的第一篇研究论文，是由一些片段组成的，分量十足却单调乏味。他未发表的作品却更具野心：他已经为方程理论作出了根本性贡献。他把自己得出的一些结果写下来，寄给了法兰西科学院，希望能在法兰西学院的期刊上发表。一位期刊方面的专家说，和现在一样，当时的所有投稿都会递给审查人，由他根据

一篇论文的创新性、价值和兴趣点进行评价。当时的审查人应该是柯西，他是当时法国的首席数学家。在加洛瓦所投论文内容所涉及的研究领域里，柯西是不二人选。

不幸的是，柯西非常忙。传闻说是柯西弄丢了加洛瓦的手稿，有人说这是因为他嫉妒加洛瓦的才华。然而，事实却没有这么复杂。柯西给法兰西学院写过一封信，日期是1830年1月18日，信中说由于自己“在家中感到不适”，所以未能提呈关于“年轻的加洛瓦”的论文的报告，信中还顺便提到了自己的著作。

这封信告诉了我们几件事。首先是柯西并没有扔掉加洛瓦的论文，而且一直保存了六个月。第二是柯西肯定读过了加洛瓦的论文，并且认为这篇论文非常重要，应该引起法兰西科学院的重视。

但是在下次会议上他只提交了自己的论文。加洛瓦的手稿弄哪儿去了呢？

法国数学家勒内·泰顿（René Taton）认为加洛瓦的论文深深震动了柯西。也可能是由于柯西太吃惊了，所以他没有给法兰西学院提交加洛瓦的论文，而是建议加洛瓦对自己的论文进行扩充，或建议他就自己的理论做进一步的研究，然后参评法兰西学院的数学大奖，这是一项非常高的荣誉。我们现在并没有关于柯西这些建议的资料证据，但是我们知道，1830年2月，加洛瓦确实向法兰西学院大奖投了稿。

我们并不太清楚加洛瓦这个文件的详细内容，但是也可以通过他的遗稿猜出个大概。如果加洛瓦那些重要的著作能够悉数保存下来，历史也许就不是我们现在看到的这个样子了。可惜，这些手稿都遗失了。

1831年，圣西门主义（一场新教社会学运动）的刊物《环球》（*Globe*）给出了一个大致可信的解释。《环球》上讲了一场官司，控

告加洛瓦公然威胁国王的人身安全。文章还说："这部论著……应该赢得大奖，因为它解决了拉格朗日未能解决的问题，柯西同意将大奖授予作者。但不知何故，这部论著丢失了，大奖也根本没有这位新人的份儿。"

我们必须弄清这篇文章的事实依据问题。1830年9月，为了躲避革命中的反智倾向，柯西出逃国外，所以，这篇文章的内容不会出自柯西之口。相反，这些内容更像是加洛瓦自己说的。加洛瓦有一个好朋友叫奥古斯特·夏维利耶，是他邀加洛瓦加入了圣西门公社。夏维利耶很可能就是这篇文章的作者，加洛瓦可能也曾参与其中，这正是他生命中最后一段时间里发生的事。如果是这样，这个故事就一定是来自加洛瓦，可能是加洛瓦的杜撰，也可能是柯西真的夸奖过他的著作。

*

让我们回到1829年。在数学研究的前沿，加洛瓦感到越来越沮丧，他渴望能获得数学界的认可，但一直未能如愿。随后，他的个人生活开始分崩离析。

堡拉瑞恩的情况也非常糟糕。加洛瓦的父亲尼古拉斯卷入了一场麻烦的政治风波，触怒了村里的牧师。牧师用下流的手段散布关于尼古拉斯亲属的恶毒言论，还在上面伪造了尼古拉斯的签名。绝望的尼古拉斯窒息自杀了。

这件事就发生在加洛瓦参加巴黎高科的入学考试的前几天。考试情况也很糟。据记载，加洛瓦曾往一个主考官的脸上扔过黑板擦。尽管这个黑板擦可能是布做的而不是木头做的，但仍引起了主考官的不

悦。1899 年，J. 波特兰（J. Bertrand）讲了当时的详细情况。考官问了一个加洛瓦从未涉足过的问题，惹得加洛瓦大发脾气。

无论怎样说，加洛瓦没能通过入学考试，只能被困在原地。他曾经非常自负地以为自己一定会通过，从未想过一旦通不过考试该怎么办，也没有去准备师范大学（École Préparatoire）的考试。这所学校如今已更名为巴黎高师，比巴黎高科的名头要响得多，但在当时，只能屈居第二。加洛瓦急急忙忙地开始复习，数学和物理都考得很出色，文学却考得乱七八糟，但到底还是通过了。1829 年底，加洛瓦获得了科学和文学学位。

我们前面提到，加洛瓦 1830 年向法兰西学院最高奖投了自己关于方程的论文。当时的秘书约瑟夫·傅里叶（Joseph Fourier）把论文带回家中，准备抽空浏览一遍。但是不幸又一次阻断了加洛瓦前进的道路：傅里叶不久就去世了，没来得及读那篇论文。更倒霉的是，这篇论文再也没有找到。但是这个奖项还有另外三个负责人：勒让德、拉克鲁瓦克斯（Sylvestre François Lacroix）和路易·潘索（Louis Poinsot）。弄丢论文的可能就是他们三个人中的一个。

可想而知，这件事把加洛瓦气得暴跳如雷。他开始相信，这是那些庸碌之辈扼杀自己的天才的阴谋。他很快就找到了替罪羊，就是暴虐无道的波旁王朝。他立志要推翻波旁王朝。

六年前，也就是 1824 年，国王查理十世继承了路易十八的王位，但是很不受欢迎。王朝的对头自由派在 1827 年大选中表现很不一般，1830 年更是百尺竿头更进一步，成了多数派。查理十世冒着逊位的危险走了一步险棋，7 月 25 日，他提议取消新闻自由。他完全不考虑人民的情绪，致使国家很快发生了叛乱，三天后，他作出了妥协，由奥尔良公爵路易·菲利普继承自己的王位。

巴黎高科（加洛瓦一直想要就读的学校）的学生在这一事件中发挥了重要作用，他们在巴黎的大街上进行了示威游行。但是在这个命运攸关的时刻，积极的废帝主义者加洛瓦在干什么呢？他和自己师范大学的同学一起被关在学校里，校长吉尼奥（M.Guigniault）选择了自保。

加洛瓦对自己被剥夺了参与历史进程的权利感到十分愤懑，他在《学校公报》（*Gazette des Écoles*）撰文，猛烈抨击了校长吉尼奥：

> 吉尼奥昨天在您的报纸上发表的文章在我看来极不合适。我原以为您一定很希望能揭穿这个人的嘴脸。
>
> 以下就是事实，46 名学生可以为此作证。
>
> 7 月 28 日早上，当若干名师范大学的学生想要加入斗争行列的时候，吉尼奥告诉他们，如若不听劝阻，他有权叫警察来维护学校秩序。这就是为什么 7 月 28 日那天来了那么多警察！
>
> 同一天，吉尼奥还以他一贯的酸溜溜的口吻告诉我们："有很多勇敢的人同时在与两方作斗争。如果你是个士兵，我不知道你会选择哪一方。你会为自由牺牲，还是为法令牺牲？"
>
> 这就是这个人的真实嘴脸，昨天，他还在自己的脑袋上扣了个三角帽（共和派的标志）。这就是我们的自由！

第二天，编辑发表了这封信，把作者的名字去掉了。由于加洛瓦发表了这篇匿名文章，他很快就被校方开除了。

加洛瓦报复性地加入了国家卫队，当了炮兵，这支队伍是培养共和主义者的温床。1830 年 12 月 21 日，这支部队被派驻在卢浮宫附近，加洛瓦很可能就在队伍里。四名行政长官被推上了法庭，公众情绪十

分激烈：他们希望将这几个人判刑，否则大家就要暴动。但是在判决公布之前，国家卫队的炮兵队就从那儿撤退了，那里被国家卫队的正规军和忠于国王的其他部队接管。宣判的结果公布了，几个官员被判监禁，暴动没有发生。数日后，路易·菲利普觉得国家卫队炮兵队对自己的安全是个威胁，就把它解散了。和自己的数学生涯一样，加洛瓦的革命生涯也没有成功。

如今，现实问题变得比政治问题更加紧迫——他必须养活自己。加洛瓦当上了私人数学教师，40个学生跟着他学习高等代数。我们知道加洛瓦的文字表达能力不太好，所以可想而知，他教得也不太好。加洛瓦可能在课上掺杂了不少政治评论，这对一般人来说无疑太难懂了。不管怎样，学生数量急剧减少。

加洛瓦仍然没有放弃数学事业，他将自己著作的第三个版本投给了法兰西学院，命名为“论用根式解方程的传统”。由于柯西已经逃离巴黎，审查人变成了西莫恩·泊松（Siméon Poisson）和拉克鲁瓦克斯。两个月过去了，加洛瓦没有收到任何回音，于是他写了一封信询问情况，但没有人给他回信。

1831年春天，加洛瓦的脾气变得更加暴躁。4月18日，数学家苏菲·热尔曼（Sophie Germain，在她刚刚开始自己的研究的1804年，就给高斯留下了深刻印象）给纪尧姆·里布里（Guillaume Libri）写信讲述加洛瓦的情况：“他们说他会完全疯掉，恐怕这是真的。”加洛瓦一直都是那种情绪不稳定的人，而现在几乎成了一个十足的偏执狂。

那一月，权力机构因为卢浮宫事件逮捕了19名炮兵队的士兵，并以煽乱罪把他们推上了法庭，但是陪审团宣判他们无罪。5月9日，差不多二百名共和主义者士兵在勃艮第葡萄酒饭店举行了庆祝宴会，每个人都希望推翻路易·菲利普的统治。小说家大仲马当时写道：“整

个巴黎都很难见到这样的事情，二百多个反对政府的人在下午五点钟聚集在花园上第一层的长厅里。”这件事动静越来越大，有人曾看见加洛瓦一只手拿着酒杯，另一只手拿着一把匕首。活动参与者认为这是在向国王示威，以此来表示自己的决心，这些人最后还鼓噪着上了大街。

第二天早上，加洛瓦在母亲的房子里被捕了（有人认为宴会上有警察局的间谍），被控威胁了国王的安全。这一次，他有了一点政治头脑，在审判上温和地承认了一切：他说自己本来很敬仰路易·菲利普，而自己挥舞匕首时加了一句限定语，“如果他变成了一个卖国贼。”他可怜巴巴地说，这句话被当时嘈杂的环境淹没了。

尽管加洛瓦言之凿凿，但是他确实希望路易·菲利普背叛法国民众。审判员问加洛瓦“是否认为国王会废止自己的法令”，加洛瓦回答说，“如果他还没有放弃，那他很快就会成为一个卖国贼。”他甚至毫不掩饰地表明了自己的看法：“如果他还没有废止那套法令，政府就有理由认为路易·菲利普迟早有一天会叛国。”尽管如此，加洛瓦还是被判无罪，也许是因为法庭也认同他的所作所为吧。

6 月 15 日，加洛瓦被释放了。三个星期后，法兰西学院评审他的论文。泊松觉得这篇论文“无法理解”，原文是这样说的：

> 为了理解加洛瓦的证明，我们想尽了办法。他的推论不太清晰，也不太成熟，因此我们无法判断其正确性，因此也不知该在这篇报告里说些什么。作者称这篇论文的目的是阐明一些数学方法中可能存在的普遍原理。也许这篇论文透露了某个普遍原理中互相纠结的各部分中的某些部分，如果把这些部分放在一起理解，应该比单独理解更容易。我们希望作者能发表完整的著作，以便阐明一个更清楚的观点。但是，对于他这次往学院投稿的论

文，我们觉得不能通过。

最倒霉的是，这篇报告在当时看来已经是很厚道了。审查人是这样说的：

> （这篇论文）空洞无物，看题目就知道，这篇论文是写用根式解方程的可能性问题的。事实上，就算加洛瓦的观点是正确的，也无法据此断定一个高次方程是否可以用根式求解。因为人们必须首先弄清楚这个方程是不是不可约方程，然后还要看它的任一根是否可以用另外两个根组成的有理分式来表示。

这一评断说到了加洛瓦论文的最精彩之处，就是判断高次方程是否可以用根式求解的标准问题。如何将这种检验方法具体运用到每个高次方程中确实不够明了，因为在进行检验之前必须知道方程的根。但是，在没有公式的前提下，怎么才能知道方程的解呢？正如泰格诺（Tignol）所说："加洛瓦的理论出乎人们的意料，它太新颖了，人们根本接受不了。"评审人希望看到决定方程可解性的"系数"条件，而加洛瓦却把方程的根当做决定条件。然而评审人这种期望是不合逻辑的，因为这里既没有简单的系数标准，也没有其他的此类标准。但是"后见之明"并没能帮上加洛瓦的忙。

*

7月14日，巴士底日，加洛瓦和自己的朋友恩内斯特·杜厦莱特（Ernest Duchâtelet）走在了共和主义者示威队伍的最前面。加洛瓦穿

着被解散了的炮兵队的制服，身上带着一把小刀、几把手枪和一杆装满了弹药的来复枪。在当时，穿这身制服和携带武器都是非法的。在新桥上，两个人都被捕了，原因是穿制服有轻微的违法嫌疑。他们被送进了圣·佩拉吉监狱（Sainte-Pélagie）等待审判。

在监狱里，杜厦莱特在自己囚室的墙上画上了国王的脑袋，以此表示要把国王推上断头台。这对于这场官司毫无帮助。

杜厦莱特首先被推上法庭，紧接着就是加洛瓦。10 月 23 日，审判结束，加洛瓦被判了刑，他的上诉也在 12 月 3 日被驳回。这时，他已经在监狱里关了四个多月了，但审判的结果是，他必须再服刑六个月。他做了一段数学研究，然后就染上了 1832 年那场霍乱，被送到医院，并被保释出狱。在这段自由的日子里，加洛瓦谈了平生第一次也是最后一次恋爱，女友是一个叫“斯蒂芬妮·D”（Stéphannie D）的女孩，加洛瓦就这样写她的名字。

围绕着这次恋爱，出现了很多史料匮乏的臆想之作。人们一度不知道斯蒂芬妮的姓氏，也不知道她是个什么样的人。这种神秘性让她成了一个浪漫的形象。加洛瓦在自己一份手稿中写出过她的全名，但后来又把它涂掉了，让人无法辨认。历史学家卡洛斯·伊恩范托奇（Carlos Infantozzi）的著作非常细致地研究了加洛瓦那份手稿，指出这位女士就是斯蒂芬妮－费里西埃·波特林·杜·莫特尔（Stéphannie-Flicie Poterin du Motel）。她的父亲让－路易斯·奥古斯特·波特林·杜·莫特尔（Jean-Louis Auguste Poterin du Motel）是西耶尔·佛尔特里埃（Sieur Faultrier）的常驻医生，加洛瓦在那儿度过了生命中的最后几个月。

不知道让－路易斯怎么看待两个人的关系，不过看起来并不像是一个一文不名、没有工作，有着极端的政治观点以及犯罪记录的危险

青年在向自己的女儿求爱。

斯蒂芬妮的观点我们倒是了解一些，不过也只是通过加洛瓦潦草地誊写的她写给他的信中的只言片语。围绕这段小插曲有很多谜团，与接下来发生的事有着至关重要的关系。很明显，加洛瓦被拒绝了，这让他十分伤心，但具体情况我们不太清楚。他是否因自己的一片痴情没有得到任何回报而耿耿于怀？斯蒂芬妮是否鼓励过他进一步去追求自己？斯蒂芬妮是否因紧张而退缩了？违抗父命的顾虑无疑让斯蒂芬妮瞻前顾后。

对加洛瓦来说，这段感情肯定十分重要。5 月，他写信给好友夏维利耶："我怎能眼睁睁地看着自己作为一个男人的所有幸福的赌注全部付诸东流呢？" 在他的一页遗稿的背面，抄写着斯蒂芬妮两份信件的片段。其中一封是这样开头的："让我们断绝这种关系吧"，这意味着他们之间确实存在着某种关系，"也不要再想着那些根本就不存在也不可能存在的事情了"，却给人一种完全相反的印象。在另一封信中有这样几句话："我听从了你的建议，也认真想过……那些……曾经……但是，先生，那也只是如此而已。你想错了，而且也不应该因此而懊悔什么。"

也许，加洛瓦已经预料到这件事的结果，他的感情是不可能得到回报的；抑或是一连串的拒绝对他来说就是某种形式的鼓励，看起来加洛瓦是陷入了最严重的单相思之中。他们的爱情是不是预示着更不祥的事情？加洛瓦与斯蒂芬妮断绝关系之后（或者说在加洛瓦自称的"断绝关系" 之后），有人提出要跟他决斗。对手提出决斗的幌子就是他反对加洛瓦追求这位女士，但这件事的真实情形也是一个谜。

这个故事的公认版本是：这是一个政治阴谋。作家埃里克·坦普尔·贝尔（Eric Temple Bell）和路易斯·科尔罗斯（Louis Kollros）认

为，加洛瓦的政敌发现了他对莫特尔小姐的迷恋，认为这是以“事关名誉”的借口剪除异己的好机会。还有一个更离奇的说法是，加洛瓦是被警察局的间谍杀害的。

这种说法现在看起来并不可信。在自己的《回忆录》中，大仲马认为加洛瓦是被彼歇·德·艾尔宾维尔（Pescheux D' Herbinville）杀死的。彼歇·德·艾尔宾维尔是一个共和主义者，大仲马曾这样描述他：“一个迷人的青年，他会用绢纸做弹药筒，并用玫瑰色绦带来装饰。”这东西就是爆竹的早期形态，如今的圣诞节上已经很常见了。从某种意义上说，德·艾尔宾维尔是贫农阶级的英雄，是因阴谋推翻政府罪被逮捕、后来又被无罪释放的19名共和主义者之一。他也肯定不是警察局的间谍，因为在1848年科西迪耶尔（Caussidiere）掌管警察局之后，把此类间谍的名字全都列了出来。

警察局关于决斗的报告称，另一位决斗者是加洛瓦的革命同志，决斗的真实情况与表面看起来的情况没什么两样。加洛瓦自己的话也证实了这种说法：“我希望爱国主义者们和我的朋友不要指责我没有为国家而死。我是因一个无名的婊子而死的。我将在痛苦的诽谤中结束我的生命。天哪！为什么死得这么微不足道，为什么要死在这么窝囊的事情上！……宽恕那些杀我的人吧，他们有着美好的信仰。”也许是他没有意识到自己是政治阴谋的牺牲品，也许根本就没有什么阴谋。

斯蒂芬妮至少是这次决斗的直接原因。在出发去决斗之前，他还在自己的桌子上留下了潦草的字迹，其中就有“一个女人”（Une femme），第二个单词又被他涂掉了。但是，决斗的最终原因和加洛瓦的很多其他的故事一样不为人知。

数学的故事却相对清晰得多。5月29日，决斗前一天，加洛瓦写

信给奥古斯特·夏维利耶，简要阐述了自己的发现。后来，夏维利耶在《百科评论》(*Revue Encyclopédique*) 上发表了这封信。这封信简要论述了群与多项方程的关系，指明了用根式解方程的充分必要条件。

加洛瓦还提到了椭圆函数和代数函数的结合问题，还有一些无法验证的神秘问题。潦草地写在页边上的"我没有时间了"造成了另外一个神话：加洛瓦在决斗前夜发狂似的写着自己的数学发现。但这句话就写在作者注解的旁边，很难让人联想到这样的画面。另外，这封信也是加洛瓦第三份被拒绝的手稿的注解，那份手稿还被泊松加上了边注。

他们是用手枪决斗的。事后有报道说他们是相距25步开的枪，但事实上可能没有那么精准。1832年6月4日的《先驱报》(*Le Precursor*) 是这样报道的：

> 巴黎，6月1日——昨天，一场令人伤心的决斗夺去了一个大有前途的青年的生命，但是他的政治行为掩盖了他过人的才智。年轻的艾瓦里斯特·加洛瓦的决斗对手是他的一个老朋友，和他一样是"人民之友"的成员，曾因一次政治审判而广为人知。据说两个人的争斗是爱情引起的。他们选择了手枪作为武器，但由于两个人是老朋友，所以不忍心直面对方，所以他们蒙上了眼睛。他们面对面直接开了枪，但只有一支手枪开了火。加洛瓦被对手的子弹射穿了，他被送进了科钦医院 (Hospital Cochin)，并在两小时后死在了那里。他只有二十二岁，而他的对手 L. D. 比他还年轻一点。

"L. D." 指的是不是彼歇·德·艾尔宾维尔 (Pescheux D' Herbinville)

呢？也许就是他。字母D容易理解，因为当时这样拼写很常见；字母L则可能是个错误。这篇文章中的细节并不可信：决斗的日期都写错了，加洛瓦死亡的时间和死时的年龄也写错了，所以，写错首字母也很有可能。

宇宙学家兼作家托尼·罗思曼（Tony Rothman）的说法则更为可信。这种描述放在杜厦莱特身上最合适，而不是艾尔宾维尔，杜厦莱特曾和加洛瓦一起被关进了新桥监狱。加洛瓦的传记作者罗贝尔·鲍内（Robert Bourgne）和让-皮埃尔·阿兹拉（Jean-Piere Azra）指出杜厦莱特的教名是“恩内斯特”（Ernest），很可能是错了，也许字母L也是错的。用罗思曼的话说：“我们可以想像一对好朋友爱上同一个女孩的可靠可信的情景，最后他们决定用可怕的俄国轮盘赌式的手段决出胜负。”

这种说法也存在着很多让人想不明白的地方。加洛瓦被击中了腹部，伤在这里几乎没有生还的可能。如果决斗是面对面射击，这就没什么奇怪了；而如果是在25步之外，那就只能证明加洛瓦的运气真是坏到了极点。

加洛瓦并不像《先驱报》所说的那样死于两个小时后，而是于第二天死在科钦医院里，也就是5月31日。他的死因是腹膜炎，并且拒绝牧师的祈祷。1832年6月2日，加洛瓦被埋葬在蒙帕纳斯公墓的一个普通墓穴中。

他写给夏维利耶的信是这样结束的：“请求雅各比或者高斯就我的理论的重要性公开发表他们的意见，而不是关于事情的真相。然后，我希望这些理论能够帮助人们解开所有难题。”

*

加洛瓦到底作出了什么成就？其中的“难题”指的又是什么呢？

这两个问题的答案就是这本书的中心问题，三言两语很难说得清楚。加洛瓦开启了数学研究的新视野，他改变了数学的内容，向抽象领域迈出了必不可少又不易理解的一步。在加洛瓦这里，数学不再研究数和形状——算数和几何，以及由二者发展出来的代数和三角法之类的思想。数学从此开始了结构问题的研究，从“事物”研究变成了“过程”研究。

我们不能把这一转变的所有功劳都给加洛瓦，因为他的成就建立在拉格朗日、柯西、鲁菲尼和阿贝尔的研究所奠定的基础上。他以高超的技巧把这些成果变成了自己的东西。他是第一个真正理解数学问题可以转换成更抽象的思想，以便更好地进行解释的人。

加洛瓦的理论的价值没过太久就引起了数学界的注意，但事实上却差一点就被人们遗漏了。是约瑟夫-路易·刘维尔拯救了它。刘维尔是拿破仑一位将军的儿子，当上了法兰西学院的教授。1843年夏天，刘维尔告诉法兰西学院（加洛瓦的三篇论文都在这儿弄丢或被拒绝了）：“我希望引起学院对加洛瓦论文的兴趣，这篇论文写得既扎实又精确，我在其中找到了一个绝妙问题的答案：一个方程是不是可以通过根式求解……”

如果不是刘维尔不辞辛劳地研读了这位倒霉的革命者潦草而混乱的手稿，并花费大量的时间和精力猜出作者的意图的话，这份手稿就和杂物一起被扔掉了，关于群论的思想只能有待于下一次的发现了。因此，很多数学家对刘维尔非常感激。

随着对加洛瓦理论的认识的深化，一种崭新而有力的数学概念出

现了，这就是群的概念。群，作为数学的一个独立分支、一种对称性运算出现了，如今，群的观念已经渗透到了数学的各个方面。

*

加洛瓦研究了排列组合 —— 重组一串对象的方法。加洛瓦研究的是代数方程的排列组合。最简单有趣的是三次方程的三个根，a，b，c。我们可以按照拉格朗日和鲁菲尼的方法回想一下这三个符号的六种排列组合，我们可以依次将其中任意两个组合相乘。举个例子，我们可以看到，$cba \times bca = acb$。如此这般，我们可以建立一个关于这六个组合的“乘法表”。如果我们把这些排列组合用符号标记出来，比如 $I=abc$，$R=acb$，$Q=bac$，$V=bca$，$U=cab$，$P=cba$，下一步就比较容易进行了。这样乘法表就会变成：

	I	*U*	*V*	*P*	*Q*	*R*
I	*I*	*U*	*V*	*P*	*Q*	*R*
U	*U*	*V*	*I*	*R*	*P*	*Q*
V	*V*	*I*	*U*	*Q*	*R*	*R*
P	*P*	*Q*	*R*	*I*	*U*	*V*
Q	*Q*	*R*	*P*	*V*	*I*	*U*
R	*R*	*P*	*Q*	*U*	*V*	*I*

三次方程根的六种排列组合的乘法表

横轴上的 X 与纵轴上的 Y 的交汇点就是 XY 的乘积，也就是说：“先算出 Y，再算出 X。”

加洛瓦发现这个表有一个简单而重要的关键特征。任意两个排列

组合的乘积本身就是一个排列组合，也就是 I，R，Q，V，U 和 P 这六个符号中的一个。更小的排列组合集有着相同的“群性”：集中任意两个排列组合的积都能在集群之内找到。加洛瓦把这种排列组合集叫做“群”。

例如，集（I，U，V）会形成一个小一点的表：

	I	U	V
I	I	U	V
U	U	V	I
V	V	I	U

包含三个排列组合的子群的乘法表

在这个表中，只出现了这三个符号。如果一个群是另一个群的一部分，我们就称之为子群。

其他的子群，（I，P）、（I，Q）和（I，R），只有两个排列组合；而它们的子群（I）则只包含排列组合 I。由此可以证明，我们所列出的六个子群就是这三个符号组成的所有排列组合的群的单子群。

因此，加洛瓦说（当然，并不是原话），如果我们选定了一个三次方程，可以检验它是不是一个对称——包含根之间的所有代数关系的排列组合。举个例子，$a+b^2=5$，这就是根 a 和 b 的代数关系。排列组合 R 是一个对称吗？如果我们要检验上面的定义，可以让 R 中的 a 保持不变，把 b 替换成 c，就要求 $a+c^2=5$ 也必须成立。如果不成立，R 就很明显不是一个对称；如果成立，再检验这些根的任一种有效的代数关系，如果 R 通过了所有的检验，它就是对称的。

准确地找到哪个排列组合是一个给定方程的对称是一项很有难度的技术性活动。但是有一点是我们不用通过运算就可以确定的。一个

方程的所有对称（名词）的集一定是由方程根的排列组合组成的群的子群。

为什么？举例来说，假设 P 和 R 包含了方程根之间的所有关系，如果 R 适用于某种关系，我们就得到一个有效关系；如果 P 也适用于这种关系，我们就又得到一个有效关系。这种关系也就适用于 PR，因此，我们就称 PR 是一组对称。也可以说，由这些对称组成的集也具有群的性质。

加洛瓦的所有研究都围绕着这个一目了然的事实。这告诉我们，每一个代数方程都有一个相关的群，也就是它的对称群，为了纪念它的发明者，人们把它叫做“加洛瓦群”。一个方程的加洛瓦群就是此方程根的所有排列组合组成的群的子群。

这种见解预示着一系列问题的突破口。我们可以确定什么样的情况下会出现什么样的子群。尤其当方程可以用根式求解时，方程的加洛瓦群的内部结构就应该符合这种情形。现在给出一个方程，你就可以通过检验其加洛瓦群是否符合相应的要求，来确定它能不能用根式求解了。

*

就这样，加洛瓦可以通过另一种视角重新思考问题了，他再也不用背着袋子和梯子建塔了，而是种了一棵树。

正如阿贝尔没有把卡尔达诺的理论叫做塔一样，加洛瓦也没有称自己的理论为树。但是，加洛瓦的理论可以画成一棵树的主干上长出来的许多树枝。主干就是方程的加洛瓦群，而这些分叉、小枝和树叶毫无疑问就是子群了。

只要我们用到根式，就要弄清方程的对称怎么变化，而要弄清这

种变化，就必须研究子群。方程的群会发生什么样的变化呢？加洛瓦演示说，如果我们要求 p 次方根，群就会分成 p 个大小相等的区块。（阿贝尔认为，我们可以假设 p 一定是质数。）那么，由 15 个排列组合组成的群可以分成五个数量为 3 的群，也可以分成三个数量为 5 的群。关键问题是，我们要保证这些区块满足某些严格的条件，特别要满足的条件是，这些区块中必须有一个自身能够组成一个被称为“指数 p 的正规子群”的子群。我们可以将其看成一棵树的主干长出的 p 个小树枝，其中的一枝就是一个正规子群。

一个群的正规子群有：包含三个符号的全部六种组合的完整群（I，R，Q，V，U，P）、子群（I，U，V）和只包含一个排列组合的子群（I）。另外三个只包含两个排列组合的子群却不是正规子群。

比如，我们要解一个一般的五次方程。这个方程有 5 个根，所以排列组合就包括 5 个符号，而这样的排列组合正好有 120 个。如果方程是完全对称的，那么它的系数就有一个包含这全部的 120 种排列组合的群。这个群就是树的主干。而每一个完全不对称的根都有一个只包含一种排列组合的群——最细小的一个。因此，这棵树有 120 片叶子。我们引入对称性结构的树干、分叉、树枝和树叶的目的是要发现某种重要性质，这种性质会在我们试图得出解方程的公式的时候出现，前提是假设方程的根是通过根式表达的。

为了讨论解方程的公式，可以首先假设我们要求一个五次方根。那么包括 120 个组合的群就被分成了五块，每块包含 24 个组合。所以，这棵树长出了五个分叉。确切地说，这些分叉必须是与指数 5 的正规子群一一对应的。

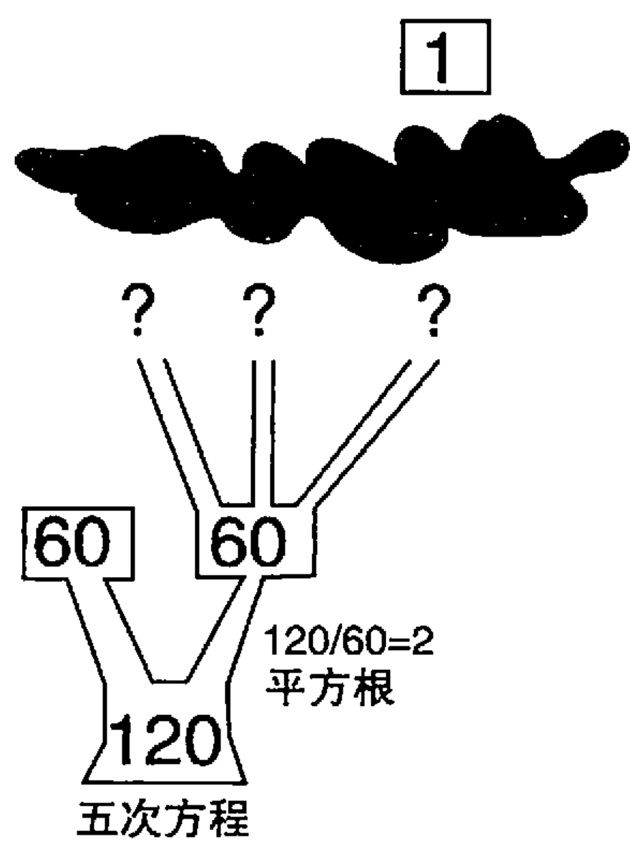
但是加洛瓦仅仅通过推算排列组合就证明：根本不存在这样的子群。

我们也可以认为这个方程有七个根。这样，120 个排列组合就必须分成相等的七块，但这是不可能的，因为 120 不能被 7 整除。事实上，除了 2，3，5 之外没有其他的原根，因为 120 只有这三个素因数，而我们又已经排除了 5。

那么存在一个三次方根吗？不巧，也没有。因为包含 120 个排列组合的群并没有指数 3 的正规子群。

现在只剩下二次方根了。一个包含 120 个排列组合的群有一个指数为 2 的正规子群吗？确实有这样一个二次方根，而且只有一个。这个子群叫交代群，包含 60 个排列组合。这样，通过加洛瓦的群论，我们就建立了一个以二次方根开始，继而得出交代群的解方程公式。现在第一个分叉处只分出了两个分叉。然而这棵树有 120 片叶子，所以必须进行再分割。这两个分叉是怎么分的呢？

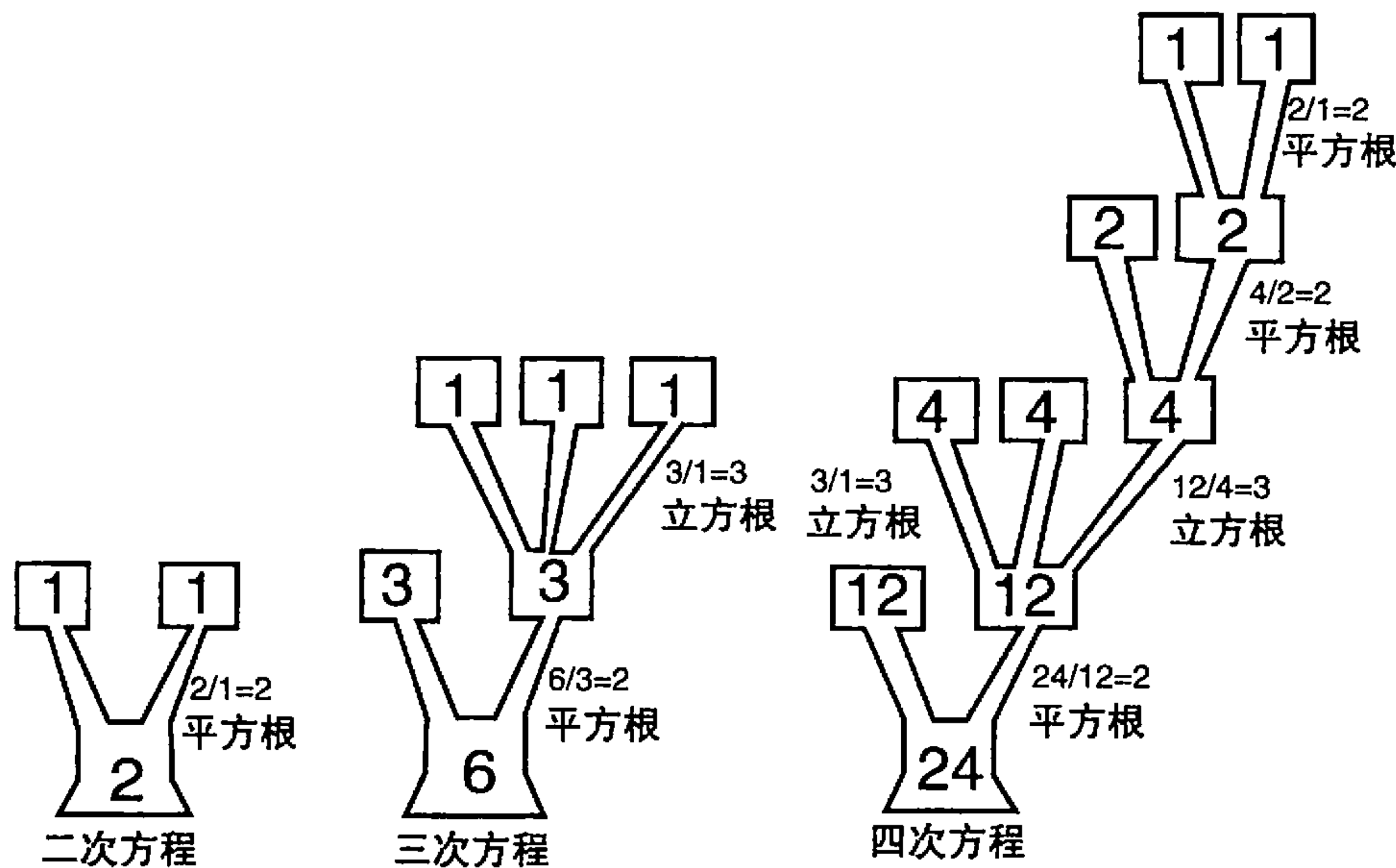
60 的素因数仍然是 2，3 和 5。所以每个分叉必须分成两个、三个或五个树枝。因此我们必须映射出相应的二次、三次或五次方根。另外，当且仅当一个交代群包含一个以 2,3 或 5 为指数的正规子群时，这一步骤才可以实现。



加洛瓦对五次方程不可解的证明

但是存在这样一个正规子群吗？这是个纯粹关于五个符号的排列组合问题。通过分析这些排列组合，加洛瓦证明：交代群根本就没有子群（除了完整群和平凡子群［I］）。交代群是“简单”的群——组成所有可构建的群的最基本元素之一。

通过逐步分成质数个分支的方式连接树干和树叶的正规子群少得可怜。因此，从第一步引入一个平方根之后，用根式解五次方程的过程就陷入了寸步难行的境地。走到这儿已经无路可走了。由主干通往每一片叶子的树不存在了，这也就意味着，用根式求五次方程的根的公式不存在了。



用群解二次、三次和四次方程

这一思想同样适用于六次、七次、八次和九次方程——所有五次以上的方程都是如此。这会让我们迷惑，为什么二次、三次和四次方程可以通过根式求解呢？事实上，群论也明确地告诉了我们怎样解二次、三次和四次方程。下面我就删去枝叶，简要地讲一下。他们和经典公式是相对应的。

下面我们就领略一下加洛瓦思想之美。它不但证明了一般的五次方程无法用根式求解，还解释了为什么一般的二次、三次和四次方程可以求解，并且粗略地告诉了我们怎么求解。在另外一些著作中，它还告诉了我们具体的求解方法。最后，他还区分了可解的五次方程和不可解的五次方程，并给出了可解方程的解法。

加洛瓦群告诉了我们想要了解的关于方程解法的所有问题。但是，为什么泊松、柯西、拉克鲁瓦克斯和其他一些专家看到加洛瓦的论文时没有任何欣喜之情呢？

加洛瓦群隐藏着一个可怕的秘密。

*

这个秘密就是：想要得出一个方程的群，就必须弄清楚方程的根的性质。但是显然，弄清方程的根的性质，我们有可能想到的方法就是知道方程的根是什么。但是别忘了，我们殚精竭虑地解方程，就是为了找到方程的根啊。

假如有人给了我们一个方程，比如 $x^5-6x+3=0$，或 $x^5+15x+12=0$，要求我们用加洛瓦的方法确定其能否通过根式求解。这个问题看起来合情合理。

但是很糟糕，加洛瓦的方法却回答不了这个问题。我们可以认为，相关的群很可能包含了 120 个排列组合——如果是这样，这个方程就无法求解。但是我们并不能确定这个群到底是不是真的有足足 120 个排列组合。也许，这五个根受着某种特殊的限制。换句话说，这项工作需要通过未知的而不是已知的条件来完成。

现在，你可以到一个数学网站，输入你的方程，网站就会帮你推

算出这个方程的加洛瓦群。我们可以了解到，上面给出的第一个方程不能通过根式求解，第二个方程却可以。并不是计算机发现了解决这一问题的方法，而是有人发现了它。在加洛瓦之后，这一领域最重大的发现就是找到了推算所有方程的加洛瓦群的方法。

加洛瓦并不具备这种技巧。找到推算加洛瓦群的可行方法又花了一个世纪的时间。由于没有这种方法，泊松和柯西有了推脱的借口。他们完全有理由抱怨说加洛瓦的思想并没有解决如何确定方程是否可以用根式求解的问题。

但是他们却没有注意到，加洛瓦的方法确定解决了一个与此大同小异的问题：方程根的哪种性质决定了其是否可解。这一问题有一个简练然而深奥的答案。这个他们希望他解决的问题……并没有一个明确的答案，并没有一个可以通过方程系数的运算来确定方程是否可解的简单方法。

*

到现在为止，我们将群解释成对称的过程中用了太多的比喻。现在我们让这个问题更明确一些，这需要更多的几何学视角。加洛瓦的继承者很快意识到加洛瓦群与对称之间的关系更容易用几何的方式理解。事实上，很多学生就是通过这种方式开始了解这一问题的。

为了获得对这种关系的感性认识，大家可以快速浏览一下我最喜欢的一个群：一个等边三角形的对称群。最后，我们会问一个很基本的问题：对称到底是什么？

在加洛瓦之前，对这一问题的所有答案都是含糊不清的，大致是指典型的具有优美的比例的东西。这种说法是无法用数学的方式来把

握的。这一问题在加洛瓦之后——在他的观念被大家接受，以及没过多久，传统的观念被摒除出数学世界之后——有了一个简单而明确答案。首先，“对称”这个词必须被解释为“一个对称”。一个对象不只有惟一的一种对称，它通常有很多种不同的对称。

那么，什么是对称呢？一个数学对象的对称是指，一种能够保存对象结构的转换。我马上就会对这个定义进行解释，但首先要明确的是，对称是一种性质而不是某样东西。加洛瓦的对称是指方程根的排列组合，而排列组合是重组事物的某种方式。严格地说，它并不是排列组合，而是应用于排列组合的法则。不是菜肴，而是烹饪方法。

这看起来就像从一端劈开一根头发一样复杂，但这对我们讨论的问题来说非常重要。

要定义一个对称，有三个关键词：“变换”（transformation）、“结构”（structure），以及“保存”（preserve）。让我以等边三角形为例来进行解释。这种三角形有三条长度相等的边和三个大小相等的角（60度）。这种性质让它很难在不同边之间进行区分，“最长的边”这种说法毫无意义，几个角之间也无法区分。我们可以说，这种无法区分各边和各角的情形是等边三角形的对称造成的。事实上，正是这些性质定义着对称。

让我们依次思考一下这三个词。

变换：我们可以对这个三角形进行一些操作。原则上讲，我们可以将它倾斜、旋转一定的角度，揉皱，像皮筋一样拉伸，染成粉红色。但是我们的选择被第二个词限制了。

结构：我们的三角形的结构是由非常重要的数学特征构成的。三角形的数学特征是指“有三条边”、“每条边都是直的”、“每条边长7.32英寸”、“在同一平面位置上”等。（在数学的其他分支中，这些

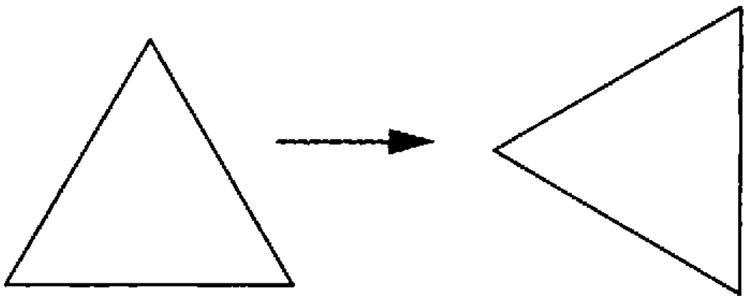
重要的数学结构会有所不同。比如，在拓扑学中，重要的是三角形要形成一个闭合回路，而它的三个角和边的直线性就不再重要了。）

保存：变换后的对象的结构必须与原来的对象进行比较。变换后的三角形必须同样有三条边，所以揉皱是不行的；边必须仍然是直的，所以不允许弯曲；每条边必须还是7.32英寸，所以拉伸也是被禁止的；还必须是在原来的位置上，所以移动十英尺也是不行的。

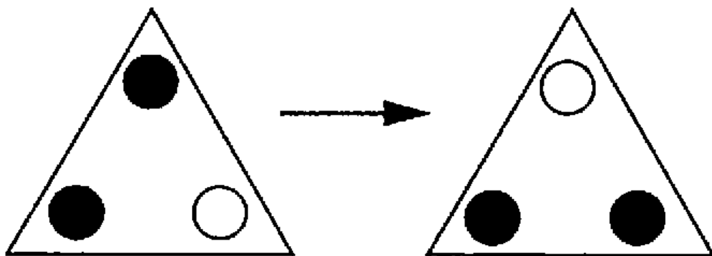
而颜色与结构没什么关系，所以染成粉红色不要紧。这样做是可以接受的，只要不影响三角形的几何性质就行。

将三角形按一定角度旋转，的确保存了某些方面的结构。如果你用硬纸板做了一个三角形，把它放在桌子上旋转一下，看起来还是个三角形。它有三条边，而且每条都是直的，长度也没变。但是三角形的平面位置还是发生了变化，这取决于你把它旋转的角度多大。

比如，如果把三角形旋转一个直角，结果看起来就不同了，各边指向的方向变了。即使在我转动三角形时你闭上了眼睛，睁开眼时还是能够发现我动了它。



旋转一个直角不是等边三角形的对称

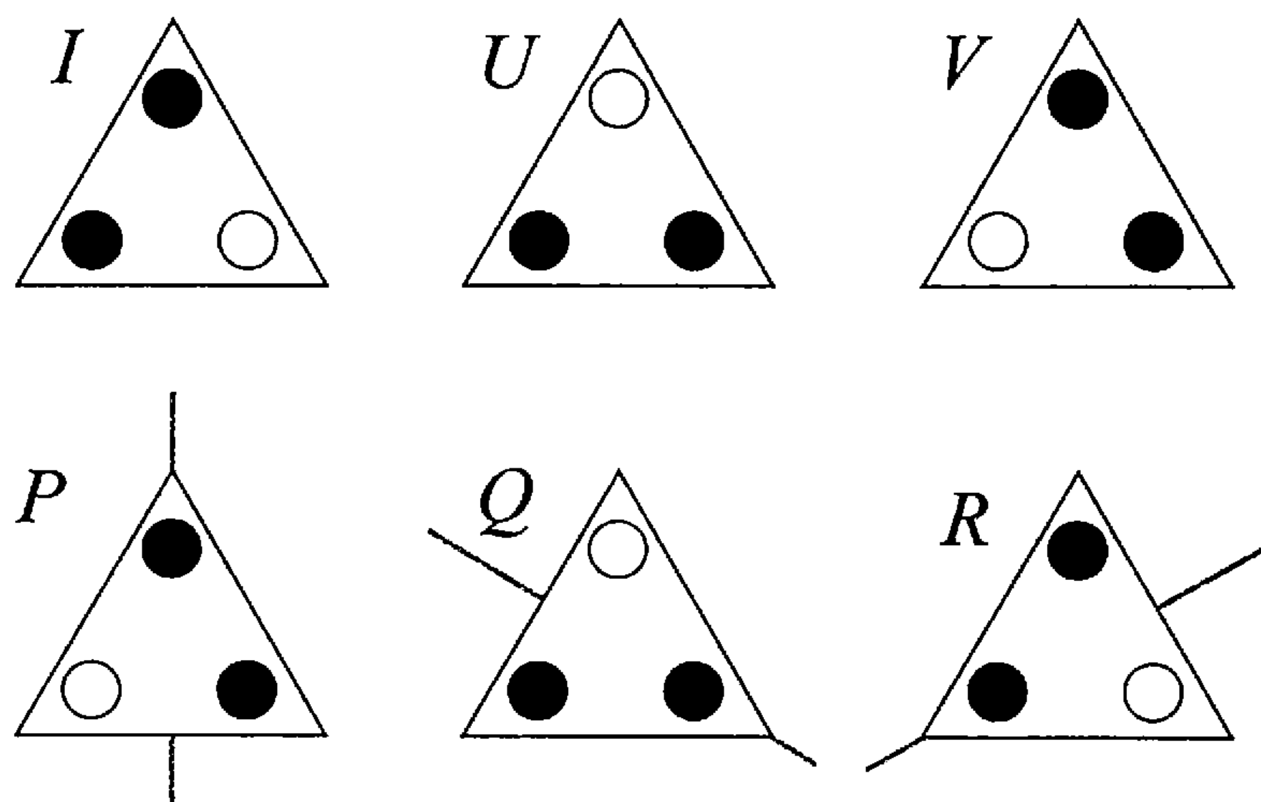


旋转120度是等边三角形的对称

如果我把这个三角形旋转120度，你就看不出前后的变化了。为了更清楚地表达我的意思，我会悄悄地在各个角标上不同的记号，这样就可以发现它往哪儿移动了。这些记号只作为标记之用，并不涉及结构的改变。如果不看这些标记，这个三角形和其他的欧几里得几何图形没什么两样，因此这个三角形和转动之前，无论形状还是位置，看起来都是一样的。

这证明，一个等边三角形，正好有六种不同的对称。还有一个是"转动240度"；另外三种是反射，一个角保持不变，其余两个角进行对调。那么，第六种对称是什么呢？那就是它本身，什么操作也不做。这看起来很普通，但是却符合对称的定义。事实上，无论我们想要保存对象的任何结构，这种变换都符合对称的定义。如果什么都不做，那就不会有任何改变。

这个普通的对称叫做"恒对称"。它看起来好像并不重要，但是少了它，运算会变得非常麻烦。这就像是做加法时少了零，做乘法时少了一。有了恒对称，一切都会变得简洁明了。



等边三角形的六种对称

对等边三角形来说，你可以把恒对称看成转动 0 度角。上图描述的是等边三角形六种对称的运用。它们是可以把你用硬纸板做成的等边三角形放入原来的轮廓中的六种方法。标记线所示的是可以得到我们想要的反影的镜子的摆放位置。

现在我要告诉你，对称是代数的一部分。我下面要做的也是所有的代数学家都会做的：把所有的东西都用符号表示出来。如上图所示，我们把这六种对称分别叫做 I，U，V，P，Q，R。恒对称是 I；另外两种旋转分别叫 U 和 V；三个反影分别是 P，Q 和 R。我前面标记三次方程的根的时候用的也是一样的符号。我这样做是有原因的，一会儿我会讲到它。

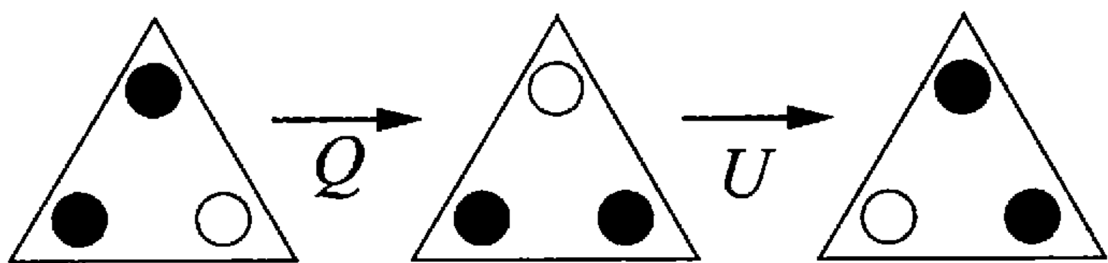
加洛瓦针对排列组合的“群性”（group property）做了大量的操作。如果你任选两个排列组合进行操作，就会得到另外一个。这对我们处理这六种对称有着巨大的启示作用。我们可以将他们任意组对相乘。可以回想一下前面的惯例：如果 X 和 Y 是对称变换，那么它们相乘所得到的 XY 就是先进行 Y 操作，再做 X 操作的结果。

假如我们想得到 VU，这就意味着我们可以先把三角形进行 U 操作，然后做 V 操作。我们可以看到，U 操作是旋转 120 度，接着的 V 操作就是在此基础上旋转 240 度。这样，UV 就是旋转 $120° + 240° = 360°$。

噢，我忘了总结一下了。

我们现在就开始总结吧。如果你把一个等边三角形旋转 360 度，一切就又都回到了原点。在群论中，这就是最终结果，而不只是进行旋转。在对称理论中，两个对对象造成相同结果的对称，都被视为是相同的。由于 UV 和恒对称有着相同的结果，因此我们可以说 $UV=I$。

再举一个例子，可以看一下 UQ 的结果？其造成的变换是这样的：



如何将对称相乘

我们会发现最终结果就是 P。因此 $UQ=P$。

这六种对称可以得到 36 种结果，其运算过程可以用一个乘法表来概括。三次方程的根的排列组合得到的也是同一个乘法表。

*

事实证明，这一明显的重合就是整个群论中最有用的技巧之一。这一技巧首次出现于法国数学家凯米尔·若当（Camille Jordan）的著作中，他令人信服地将群论变成了一个独立的学科，而不再仅仅是一种分析方程是否可以用根式求解的方法。

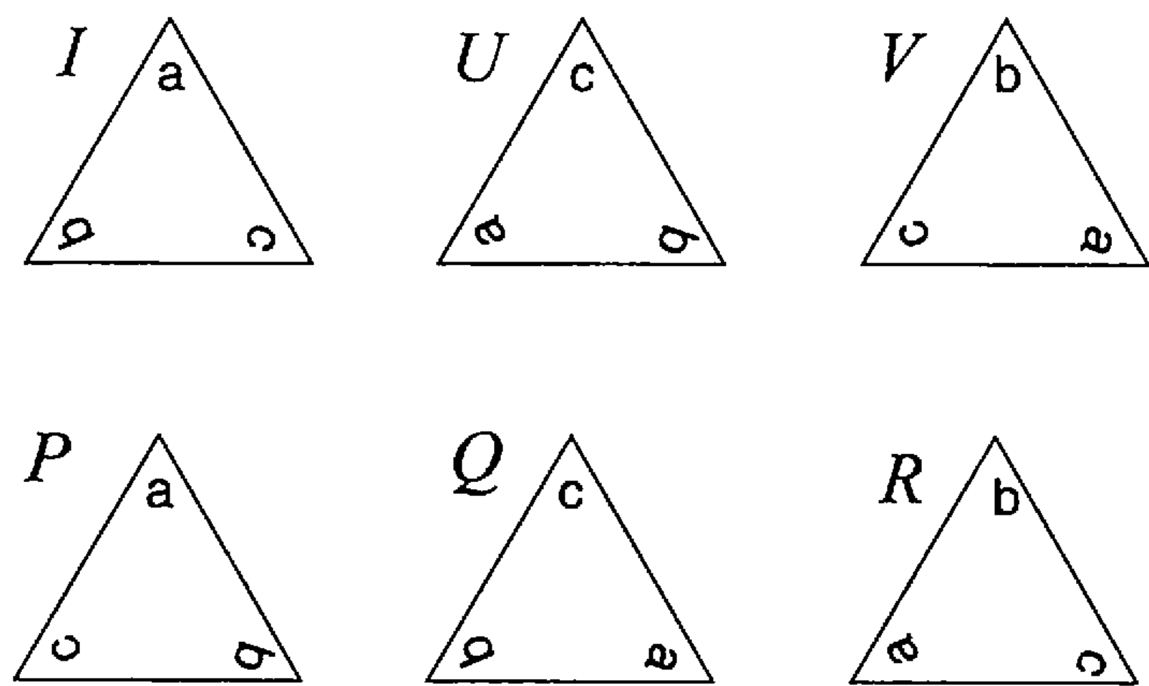
1870 年前后，若当开始了当今被称为“表示论”（representation theory）的问题的研究。对加洛瓦来说，群是由排列组合组成的——一种混排符号的方法。若当开始考虑混排更多、更复杂的空间。在数学中，最基本的空间就是乘法空间，其最重要的特征就是直线的存在。这种空间的本质法则就是“让直线是直的”，不可弯折，不可扭曲。有很多种这样的变换——旋转、反射以及改变尺寸，这些都叫“线性”变换。

英国的律师数学家阿瑟·凯莱（Arthur Cayley）发现，任何的线性变换都有一个相关的矩阵——一个方形的数字表。比如，任何三维空间的线性变换，都可以具体化为一个三行三列的实数表。因此，变

换也可以引入代数计算。

表示论会让你接触到由非线性变换组成的群。把群转换成一组矩阵的好处是，矩阵代数非常深刻而强大，若当第一个发现了这一点。

让我们以若当的眼光审视一下三角形的对称。我们不再在三角形的各角做记号，而是将其换成 a，b，c，用以代表一般三次方程的根。顺理成章地，三角形的每个对称也都被换成了这几个符号。比如，旋转 U 就是从 abc 到 cab。



等边三角形的对称和排列组合的对应

三角形的六个对称也就自然地对应于方程根 a，b，c。另外，两个对称所得的结果也分别对应于方程根的排列组合。但是平面上的旋转和反射是线性变换——它们保存了直线。所以我们也就重新解释了这些排列组合的群，也就是重新表示它，把它看成一个线性变换的群，或者说看成一个矩阵群。这一思想对数学和物理学都产生了深远的影响。

平庸的工程师和卓越的教授

对称不再是对某种规则性的模糊印象与一种对美和优雅的艺术感觉了，而是变成了具有严格逻辑定义的明确的数学观念。你可以对对称进行运算并证明关于它的定理。一个新的学科诞生了，这就是群论。这是人类探寻对称问题的转折点。这一进步为人们进行更概念化的思考提供了通行证。群的概念是从传统的数或几何图形之类的原始材料中抽出来的一个或几个抽象步骤。

通过解决古老的五次方程难题，群论已经证明了自己的价值。很快，解决其他一些古老问题的思路也变得清晰起来。你并不一定会用到群论，但是，你应该像阿贝尔、加洛瓦及他们的继承者那样思考。就算你觉得自己用不着群论的时候，它也发挥着潜在的作用。

*

在没有解决的问题中，包括古希腊几何数学家遗留给后人的三个

著名的棘手问题：三等分角、倍立方，以及化圆为方。甚至直到现在，三等分角和化圆为方仍深深吸引着众多的数学爱好者，他们无法理解为什么数学家认为解决这几个问题是“不可能的”。而倍立方却没有这么大的诱惑力。

这三个问题常常被称为“三大古老问题”，而这个短语的含义却一再扩展着。它几乎和费马最后定理一样，变成了重大的历史谜题，费马最后定理一直过了350年也没能解决。费马最后定理谜题被人们确定为无法解决的问题，而且不难确定这种观点第一次出现在哪部数学文献中。所有的数学家都清楚这一问题及其假定答案，而且清楚地了解是谁第一个解决了这一问题。

这些古希腊问题却与此不同。在欧几里得列出的未解决的问题中，并没有这几个问题。他们以缺席的方式存在着：它们是某些积极结果的明显延伸，但不知什么原因，欧几里得避开了这些问题。为什么会出现这种现象呢？因为没人能够解决这些问题。古希腊人也认为这些问题是无法解决的吗？如果是这样，也就没有人提出异议。毫无疑问，人们无法通过尺规作图法解决这些问题，就连阿基米德也不能，因为他发明了另一种技巧，但没有关于阿基米德重视是否可用图形来解决问题的证据。

后来，用图形解决问题变得越来越重要了。这些问题的解决方法的欠缺变成了人们理解几何与代数问题的重大障碍。通过某种文化渗透效应，它们变成了学者们所说的“民俗学”问题。到这些问题得到解决的时候，已经被套上了一道历史和数学意味的光环。解决这些问题被视为一项重大的突破，特别是化圆为方问题。但是对于这三个问题，答案是一样的：“无法做到。”这些问题用传统的尺规作图法是无法解决的。

这确实是一个让人丧气的答案。人一生的大部分时间都在试图找到解决问题或克服困难的方法。如果一座高建筑无法用砖和泥沙来建造，工程师会使用钢架结构和精炼水泥来建造。但是，没有人会因为指出砖不能用来建造这座高建筑而出名。

在数学上却并非如此。了解工具的局限性常常和解决问题一样重要。数学问题的重要性并不取决于它的答案是什么，而在于为什么这个答案是正确的。对于这三个古老的问题，也是如此。

*

这个证明不可能三等分任一角的人 1814 年生于巴黎，名叫皮埃尔·劳伦特·旺策尔（Pierre Laurent Wantzel）。他的父亲起初是个军官，后来成了高等商业学校（École Special du Comerce）的聘任数学教授。皮埃尔很早熟，一个熟悉旺策尔、名叫阿德玛德·让·克劳德·巴莱·德·圣-韦南特（Adhémard Jean Claude de Saint-Venant）的人，说这个男孩显示出了“惊人的数学天赋，他深深迷恋着这门学科。他甚至很快就超越了自己的导师，在他遇到了一个困难的测绘问题时，竟然派人去请才 9 岁的旺策尔”。

1828 年，皮埃尔成功地申请进入了查理曼学院（Collége Charlemagne）。1831 年，他获得了法语和拉丁语两项奖励，同时，他在巴黎高科和作为现在的巴黎高师前身的学校的入学考试中都取得了第一名的成绩，这是史无前例的。他的兴趣几乎无所不包——数学、音乐、哲学、历史——他对这些东西的爱好难分伯仲，真是一件麻烦事。

1834 年，他转向了工程学，进入了桥路学院（École des Ponts et Chaussées）。不久他就对自己的朋友坦承说自己会成为一名工程师，

“但只是一名平庸的工程师”，他说自己实际上想去教数学，并且想休一段假。转机出现了：1838 年，他变成了巴黎高科的一名分析数学讲师，1841 年，他又成了自己原来的工程学校的应用数学教授。圣－韦南特告诉我们，皮埃尔“通常在晚上工作，总是很晚才上床，然后开始读书，然后极不安稳地睡上几个小时，在这古怪又不规律的几个小时中，有时他还滥用咖啡和鸦片，吃便餐，直到结了婚才有所改变”。他娶了自己以前的拉丁语教师的女儿。

旺策尔研究鲁菲尼、阿贝尔、加洛瓦和高斯的著作，培养了对于方程问题的浓厚兴趣。1837 年，他的论文《论确定几何问题能否通过尺规作图法解决的方法》（On the means of ascertaining whether a geometric problem can be solved with straightege and compass）发表在刘维尔主编的《理论与应用数学学报》（*Journal de Mathématiques Pures et Appliquées*）上。他重新拾起了高斯丢下的图形可作图性问题。旺策尔于 1848 年去世，时年 33 岁，这可能是繁重的教学和管理事务导致的过度劳累造成的。

*

在三等分角和倍立方问题上，旺策尔的不可能证明与高斯关于正多边形的划时代研究相似，只是要容易得多。让我们从较容易理解的倍立方问题开始。我们可以用尺规作出一个边长为$\sqrt[3]{2}$的图形吗？

高斯对正多边形的分析的思想基础是，所有的几何作图都可以归结为一系列二次方程问题。在高斯看来，这是理所当然的，因为他是以代数的方式理解线与面的性质的。在很简单的代数中，任何关于可作图性的“极小多项式”——其所能够满足的最简单方程——的次数都是 2 的

幂。这个方程可能是一次方程、二次方程、四次方程、八次方程、16 次方程、32 次方程、64 次方程……但是无论次数是几，都是 2 的幂。

回过头来说，$\sqrt[3]{2}$ 满足三次方程 $x^3-2=0$，而且也是一个极小多项式。这个方程的次数是 3，不是 2 的幂。因此，若假设立方体可以通过尺规作图法放大一倍，又要在逻辑上无懈可击，那么 3 就应该是 2 的幂。很明显不是那么回事。因此，通过归谬法可以证明，这样的作图是无法实现的。

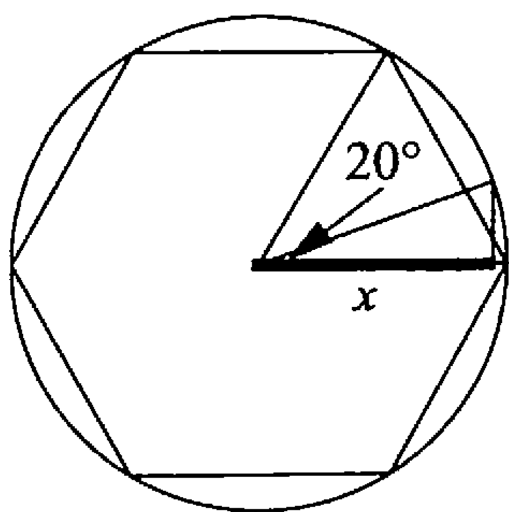
*

不可能将一个角三等分的原因与此类似，但是证明过程中牵涉到的东西稍微复杂一些。

首先，有些角可以精确地三等分。180 度角是个很好的例子，它可以分成三个 60 度的角，我们可以通过作一个正六边形来完成。所以这一证明是选择那些不能三等分的角进行的。在这些角中，最简单的就是 60 度角。60 度的三分之一是 20 度，我下面就来说明 20 度角无法用尺规作图法作出。

这一思考是发人深省的。我们可以看一下量角器这个测量角度的工具，上面明确地标着 10 度、20 度等刻度。但这些角度并不精确，最起码，刻度线是有粗细的。我们可以为建筑图或工程图画出一个合格的 20 度角。但是我们无法用欧几里得的方法画出一个完美的 20 度角，我们要证明的，就是这一论断。

解决这一难题的关键是三角学——关于角的度数的研究。让我们从一个在以 1 为半径的圆内的六边形开始。这样，我们就会找到一个 60度角，如果我们可以将其三等分，我们就可以在图形中画出一条线。



三等分 60 度角就等于作出那条长度为 x 的线段

假设这条线段的长度是 x。根据三角学，我们可以得知 x 满足方程 $8x^3-6x-1=0$。和在倍立方问题中一样，这里也是个三次方程，而且是 x 的极小多项式。但是，如果 x 是可作出的，那么它的极小多项式必须是 2 的幂。矛盾是一样的，结论也是一样的：这一假设的作图不可能实现。

我给出的这种证明方式掩盖了这一问题的深层结构，从更抽象的角度看，旺策尔解决这两个问题的方法都可以归结为关于对称的问题：与其几何性质相对应的加洛瓦群的结构不符合尺规作图法的要求。旺策尔对加洛瓦群非常熟悉，1845 年，他曾提出了方程可不可以用根式求解的另一种证明方法。这一证明紧紧追随着鲁菲尼和阿贝尔的思想，并且把二者的思想变得更加简明了。旺策尔在引言中说：

> 尽管（阿贝尔的）证明是正确的，但其形式太过复杂，而且暧昧不明，很难被普遍接受。很多年前，鲁菲尼……曾经用一种更不明晰的方式处理过同一问题……通过反思这两位数学家的研究……我们已经找到了证明这一问题的一种精确形式，并借以祛除了这一重要方程理论的所有疑问。

现在，三大古老问题只剩下化圆为方了，它可以归结为画一条长

度与 π 完全一致的线段的问题。证明这一作图的可能性比之前的问题要难得多。为什么呢？因为，π 的问题不在于极小多项式的次数不符合要求，而是根本就没有这样一个极小多项式。没有一个系数为有理数的多项式的根等于 π，你可以无限趋近 π，但不可能完全等于 π。

19 世纪的数学家已经发现有理数和无理数的区别是可以视情况而进行修饰的。无理数分为很多种。像 $\sqrt{2}$ 这种较“温和”的无理数不能用分数准确地表示为有理数，但是可以通过有理数来表示。它满足以有理数为系数的方程，$x^2-2=0$。我们可以把这种数叫做“代数的”数。

但是，数学家们意识到，应该还存在着一种“非代数的”无理数，这种无理数比那些“代数的”无理数与有理数的关系更远。它已经完全超出了有理数王国的国界。

第一个问题是，这种“超越的”数真的存在吗？古希腊人曾设想所有的数都是有理数，但希帕索斯打破了这种幻象，毕达哥拉斯勃然大怒，封杀了这一发现（希波索斯很可能就此被赶出了毕达哥拉斯教团）。19 世纪的数学家意识到，把所有的数都看成是“代数的”也有可能酿成悲剧，但是很多年来，都没有出现希波索斯这样的人物。结果，他们只能去转而求证一些实数是“非代数的”，而 π 正是一个合理的选择。但是要证明一个数（比如说 π）是无理数非常困难，你必须证明，任何一对整数相除都不等于 π。而要证明一个数是“非代数的”，你就得尝试这两个数可以组成的所有次数的方程，然后还要导出一个矛盾。这真是太难了。

第一个重大进展是德国数学家兼天文学家约翰·兰贝特（Johann Lambert）在 1768 年取得的。他在一篇关于超越数的文章中证明了 π 是无理数，而且，他的方法也为接下来的各项研究铺平了道路。他巧妙地运用了微积分，尤其是“积分”概念（任何函数的积分都是一个

函数，这个函数的变率会导致原函数的变化）。兰贝特首先假设 π 完全等于一个分数，然后开始推算自己专门为这一目的而发明的积分，这个积分相当复杂，不但涉及了多项式，还牵扯到三角函数。推算这一微积分有两种截然不同的方法，其中一种方法得到的答案是零，而另一种方法得到的答案却不是零。

如果 π 不是分数，那么这两种方法就都不适合它，所以也就不会产生什么问题了。但是如果 π 是分数，零就不等于它自己，而这是不可能的。

兰贝特证明的步骤细节很专业化，但是其原理是很容易理解的。他必须把 π 与某种简单的东西联系起来，于是他找到了三角学。接下来就是把这些东西组装起来，这样，如果 π 是无理数，那么就会出现某种特定的结果。这时候，多项式码出现了，随之而来的，还有形成积分的奇思妙想。然后，证明就变成了比较两种方法所算出的积分的问题，然后拿出两种方法各自得到的答案就可以了。这里的多项式码复杂而专业，但是对专家来说，却只是家常便饭。

兰贝特证明迈出了重要一步，但是有很多的无理数是可以用有理数解释的，其中最明显的就是 $\sqrt{2}$，一个单位平方的对角线。所以证明 π 是无理数并不能证明它不能用有理数进行表示。这只意味着人们从此不用再去寻找等于 π 的分数了，但这已经完全是另一码事了。

*

这时，数学家们面临着一个非同寻常的困境。他们已经对代数数和超越数进行了区分，并且很重视这种区分。但他们仍然无法确定超越数是否真的存在。从实际意义上看，这种假定的区分很可能是毫无

意义的。

1844年，人们才证明了超越数的存在。取得这一突破的是刘维尔，他曾在法兰西学院的废纸堆中拯救了加洛瓦的著作。现在刘维尔开始着手发明一个超越数。他给出的这个超越数是：

0.110001000000000000000001000…

一串比一串更长的0被1隔开。最重要的是，每个区间之中的0的数量必须增长得特别快。

这种数“近似于”有理数，它有一个有理数近似值——这要归功于区间中的那些0。在上述数字的较长区间中，连续出现了17个0，这就意味着前面的0.110001就是刘维尔给出的数的很好的近似值，很可能比你随便想到的一个小数要近似得多。而且，0.110001和其他有限小数一样是有理数，它等于110001/1000000。而这个数并不只是精确到小数点后第六位，而是精确到小数点后第二十三位。而下一个非零数位是第二十四位，在这个数位上是1。

刘维尔意识到，有些本来不是有理数的代数数，却总是被糟糕地变成一个近似的有理数。这不只是说这些数是无理数，而且，要获得一个很近似的有理数，就必须在分数中用到非常大的数。刘维尔有意将其称为具有很好的有理数近似值的数，比称其为代数数好多了。因此，这种数必是超越数无疑。

对刘维尔这一聪明的想法，我们能直陈的诟病就是，他的数是自己生造的，无法确定它和数学有任何联系。看起来，这些数只不过是以具有有理数近似值为由头而生造出来的。如果不是因为它可以证明超越数存在，人们根本就不用在意这种数。然而也正因为如此，人们

现在确定了超越数是存在的。

超越数是否存在是另一个问题，但超越数理论至少自有其特定的内涵。下面的任务就是证明这一理论的有趣内涵。首先，π是超越数吗？如果它是超越数，也就找到了化圆为方的古老问题的要害。所有可做图表示的数都是代数数，所以超越数就是不可作图表示的。如果π是超越数，化圆为方就是不可能的。

*

π之所以如此著名，是因为它和圆形和球体密切相关。此外，数学中还有一种特征鲜明的数，被称为e，它可能比π还要重要。这个数的值近似于2.71828，并且和π一样是无理数。这个数产生于1618年，那时对数理论才刚刚兴起。e通常用来计算在一段时间内把复利包括在内的正确利率。在1690年莱布尼茨写给惠更斯（Huygens）的信中，表示这个数的字母是b。e是1727年欧拉开始使用的，并在他1736年出版的《力学》（*Mechanics*）中首次出现。

通过复数，欧拉发现了e和π之间的重要关系。关于这一关系的公式常被称为数学中最美的公式。他证明了$e^{\pi i}=-1$（这一方程有很多直观的解释，但也牵涉到很多方程）。在刘维尔发现超越数的29年之后，证明π是超越数的另一重大突破出现了。1873年，法国数学家夏尔·埃尔米特（Charles Hermite）证明了e是超越数。埃尔米特的数学研究与加洛瓦几乎是平行的——他进入了路易大帝高中，师从理查德，试图证明五次方程不可解，也想考入巴黎高科。和加洛瓦不同，他勉勉强强地通过了考试。

埃尔米特的一个学生，著名数学家庞加莱观察过他思考问题的独

特方式："应该叫埃尔米特逻辑学家！对我来说，这是雷打不动的事实。很多方法都是以某种神秘的方式在他头脑中产生的。"在埃尔米特证明 e 是超越数的过程中，创造力发挥了非常重要的作用。这一证明也用到了微积分，他用两种方法求整数的值，如果 e 是代数数，那么两种方法得到的结果应该是不同的：一个等于零，一个是非零，其中最难的步骤是找到要求值的恰当的整数。

实际的证明大约只有两页纸，但这是多么了不起的两页纸啊！因为你可能一生都找不到合适的整数。

e 至少是一个"天然的"数学研究对象。它的突然出现无疑对复分析和微分方程理论至关重要。尽管埃尔米特没有解决 π 的问题，但他在刘维尔那个生造的例子的基础上迈出了重要一步。现在的数学家们都清楚，数学运算完全可以抛开那些看似是代数数的超越数。没过多久，艾尔米特思想的一位继承者就证明了 π 是超越数。

*

卡尔·路易·斐迪南德·冯·林德曼（Carl Louis Ferdinand von Lindeman）生于 1852 年，父亲斐迪南德·林德曼是个语言学教师，母亲艾米丽·克鲁修斯是校长的女儿。斐迪南德后来换了工作，到一家煤气厂当了主管。

和很多 19 世纪末的德国学生一样，小林德曼在不同的大学之间来回转移——哥廷根、爱尔兰根、慕尼黑。在爱尔兰根，他在菲利克斯·克雷恩（Felix Klein）的指导下完成了关于非欧几何的博士论文，取得了学位。他又旅行去了牛津大学和剑桥大学，后来又到了巴黎，并在那儿结识了埃尔米特。1879 年，林德曼在弗赖堡大学谋得了一个

教职。四年后，他又转去了柯尼斯堡大学（Königsberg），在那儿遇到了一个女演员的女儿伊丽莎白·库丝纳尔，并与她结了婚。之后的十年，他担任了慕尼黑大学的全职教授。

1882年，在柯尼斯堡的任期和去巴黎的行程之间，林德曼提出了通过扩展埃尔米特的方法来证明 π 是超越数的想法，并因此而出名。有些历史学家认为林德曼不过是走运而已——他只是误打误撞地拓展了埃尔米特的伟大思想。就像高尔夫球手加利·普莱尔（Gary Player）所说："我打得越好，运气就越好。" 对林德曼来说，很可能也是这样的情况。如果每个人都可能撞大运，那为什么埃尔米特没撞上呢?

后来，林德曼转向了数学物理领域，开始研究电子。他最有名的研究生，就是大卫·希尔伯特（David Hilbert）。

林德曼证明 π 是超越数的方法是兰贝特创造的，埃尔米特又把它向前推进了一步：写下一个合适的整数，用两种方法进行运算，如果 π 是代数数，答案就不一致。林德曼选用的整数和埃尔米特选用的整数有着紧密的关系，但可能更复杂。欧拉发现的 e 和 π 之间的关系的确是一种很美的关系。如果 π 是代数数，那么 e 就会具有某种更新奇的性质——与代数数相似，但不是代数数。林德曼证明的核心是 e，而不是 π。

林德曼的证明翻开了数学上的一个重要篇章。化圆为方的不可能性从此成了一个不起眼的小龙套，更重要的是，数学家们知道了这为什么是不可能的。现在，他们可以沿着超越数理论这条路继续前进，这是一个当今仍很有活力、也极其艰深的研究领域。至今，很多最明显、最具合理性的超越数猜想仍然没有答案。

*

我们可以通过阿贝尔和加洛瓦的思想回顾一下正多边形的作图问题。当正 n 边形中的 n 等于几时，才可以用尺规来把它作出来呢？这个问题的答案意味深长。

在《算数研究》中，高斯给出了边数 n 的充分必要条件，但他只证明了这些条件的充分性。他声称自己也证明了这一条件的必要性，但这些东西和他的很多研究成果一样，都没有发表。高斯解决了主要问题，但其细节，是由旺策尔在 1837 年填充进去的。

在讲高斯的答案之前，我们先来看一下正十七边形。数字 17 的哪些性质让正十七边形可以用尺规作出呢？ 11 和 13 为什么不可以呢。

我们可以发现，这三个数都是质数。大家很容易明白，如果正 n 边形可以用尺规作出，而质数 p 又可以整除 n，那么正 p 边形也可以用尺规作出。例如，一个正十五边形，每三个点取一个第三点，就会形成一个正五边形。所以，对于边数为质数的正多边形应该也是如此，而要用这一结论去求得一个最终的解决之道也是合乎逻辑的。

17 是一个质数，所以从它开始着手研究，是个不错的选择。我们用更现代的术语来表述的话，高斯的分析的基础就是：方程 $x^{17}-1=0$ 的解就是正十七边形的顶点。这个方程有一个很明显的解，$x=1$。另外 16 个解是一个 16 次方程的解，这个方程可以写作 $x^{16}+x^{15}+x^{14}+\cdots\cdots x^2+x+1=0$。这个正十七边形是通过解一连串的二次方程构成的，事实证明，这一问题是可以解决的，因为 16 是 2 的幂，等于 2^4。

更普遍地，由这一论断也可以证明 p 是一个奇质数，当且仅当 $p-1$ 是 2 的一个幂时，正 p 边形才是可作图的。这些奇质数被称为费

马质数，因为费马第一个对其进行了研究。古希腊人知道如何作出正三边形和正五边形。3−1=2 和 5−1=4，都是 2 的幂。所以古希腊的结论和高斯是一致的，所以 3 和 5 就是最早的两个费马质数。但是，7−1=6，不是 2 的幂，所以正七边形不可作图。

在进一步研究之后，高斯得出了结论：当且仅当 n 是 2 的幂，或是 2 的幂与明确的费马质数的乘积时，正 n 边形才是可作图的。

这就产生了一个问题——什么是费马质数？紧接着 3 和 5 的费马质数是高斯发现的，这就是 17。接下来就是 257，再接下来这个数就很大了，65537。对于费马质数，我们就知道这么区区几个。人们并没能证明，除了这几个以外还有其他费马质数存在，但是也没能证明其不存在。我们只知道，可能存在着一个数值极大的尚不为人知的费马质数。当今学界有一个流行的观点，认为这个数不会小于 $2^{33554432}+1$，而且这个数很可能就是下一个费马质数。（因为 33554432 本身是 2 的一个幂，即 2^{25}，而所有的费马质数都比 2 的幂大 1，而 2 的幂的次数也都是以 2 的幂的形式增长的。）这个数已经超过了千万进位。从高斯的伟大发现至今，我们都无法确切地知道哪个正多边形是可作图的，那些可能存在的极大的费马质数仍然是我们的知识空白。

尽管高斯证明了正十七边形是可作图的，但是他并没有描述它是如何构成的，而是用一大串数字写下了其要点：

$$\frac{1}{16}\left[-1+\sqrt{17}+\sqrt{34-2\sqrt{17}}+\sqrt{68+12\sqrt{17}-16\sqrt{34+2\sqrt{17}}-2(1-\sqrt{17})(\sqrt{34-2\sqrt{17}})}\right]$$

这个公式决定着高斯作出正十七边形的方法

由于平方根总是可作图的，相应的作图方法就暗含在这些给出的数字中。第一个明确的作图是由乌尔里希·冯·修格尼恩（Ulrich von Huguenin）1803 年完成的，1893 年，里士满（H. W. Richmond）发现了一种更简单的作图法。

1832 年，黎仄罗（Richelot）发表了一系列论文，以说明正 257 边形的作法，论文的总题目是“De resolutione algebraica aequationis $x^{257}=1$, sive de divisione circuli per bisectionem anguli septies repetitam in partes 257 inter se aequales commentatio coronata”（拉丁文：代数方程 $x^{257}=1$ 的解法，或通过反复对分角而将圆 257 等分），这个题目比他所作的多边形的边数还惹人注目。

有一个不太可信的故事，说是有一个狂热的博士生以 65537 边形的作法为论文课题，20 年后，他才带着自己的论文重新出现。事情的真相简直离奇古怪：林根大学的学生赫尔麦斯（J. Hermes）花了十年工夫，在 1894 年完成了这一作图，这一未发表的著作保存在哥廷根大学。但是很不幸，也许约翰·豪顿·康维（John Horton Conway）是惟一看过这一著作的数学家，但他对这部著作的正确性提出了质疑。

酒醉的破坏者

威廉·R. 汉密尔顿（William Rowan Hamilton）是爱尔兰有史以来最伟大的数学家。他出生于1805年8月3日和4日间的午夜。在他的一生中，他几乎从没弄清楚究竟哪天是他的生日。大多数时候，他把3日作为生日来过，但是在他的墓碑上却铭刻着他的生日是8月4日。据说，因为感情原因，晚年的时候，他把自己的生日改定在8月4日。汉密尔顿是一个非常聪明的语言学家，一个数学天才，同时也是一个嗜酒如命的人。当年他宣称要创造三元代数，但是转瞬间，他改变了初衷，决定改为创造四元代数。也正是他，永远地改变了人们对代数、空间和时间的看法。

威廉出生在一个富有的爱尔兰家庭，在兄妹四人中排行老三，两个哥哥，妹妹名叫伊莱扎。他的爸爸阿奇博尔德·汉密尔顿（Archibald Hamilton）是一个颇有经济头脑的律师。他偏爱奇数或者是数字三，这使得小威廉一度受到宠爱，但是随着时间的流逝这种宠爱日益变淡了。阿奇博尔德口齿伶俐，聪明并且拥有自己的信仰。他

把自己的鲜明个性、酒精和拥有的一切都传给了他的小儿子。威廉的妈妈，莎拉·哈顿（Sarah Hutton）来自爱尔兰的一个知识分子家庭，甚至比她的丈夫还要聪明。不过他对小威廉的影响除通过基因遗传之外则被减弱了很多，因为她的丈夫在威廉3岁的时候就把他送到了他的叔叔詹姆斯那里接受教育。詹姆斯是一个严谨并多才多艺的语言学家，他的兴趣也在很大程度上决定了威廉的受教育方向。

威廉所受教育的结果很让人惊喜，不过也未免显得局限。威廉在5岁的时候，就可以熟练地说希腊语、拉丁语和希伯来语。8岁的时候，可以说法语和意大利语，两年后，又学会了阿拉伯语和梵语；随后，又陆续学会了波斯语、叙利亚语、印度语、马来语和孟加拉语。詹姆斯本来还打算教他汉语，可是因为没有找到合适的教材而不得不放弃。他抱怨说，“小家伙在伦敦花掉了我很多钱，不过我希望这钱没有白花”。

埃利克·坦普尔·贝尔（Eric Temple Bell），一个数学家和准历史学家诘问到，“学这么多语言有什么用？”

当威廉结识美国的计算神童科尔伯恩之后，他从学习更多的世界语言中抽身出来，这对科学和数学来说是件幸事。科尔伯恩是那种奇人，就好像人们兜里的计算器一样。他拥有可以快速精确地进行计算的罕见天赋。如果你问他，1860867的立方根是多少，他几乎可以立即回答说是“123”。

不过这种计算天赋和数学能力还是存在差异的，就好像擅长记单词并不会使人成为优秀的小说家一样。除了高斯在自己的笔记和手稿里有大量的计算之外，几乎没有伟大的数学家精于计算。其他长于计算的数学家——当然，这在当时是必须的——也不会比一个合格的会计算得更好。甚至今天，计算机也并没有完全替代纸和笔的计算，或者心算，或者其他陈旧的算法，你总是可以从动手计算和推导中获得

洞见。但是如果有合适的由数学家编写的软件，任何人都可以通过一个小时的训练来快速解决科尔伯恩能做到的事情。

不过即使拥有这些，你也远不可能达到高斯的成就。

尽管科尔伯恩知道记忆能力非常重要，但他并不完全理解他所应用的技巧和诀窍。他被介绍给汉密尔顿，是希望这个小天才能够解释这些神秘的技巧。威廉不但做到了，甚至还改善并提高了一些技巧。直到科尔伯恩离开后，汉密尔顿才终于为他的高超智商找到了一个值得追求的科目。

17 岁的时候，汉密尔顿已经阅读了大量由数学天才写就的著作，并且知晓足够的数理天文知识来计算日食和月食。虽然他花在文学名著上的时间仍比花在数学上的时间要多，不过数学已经成了他的真正热情所在。不久他就取得了新的发现。就像 19 岁的高斯发现了构造正十七边形的方法一样，17 岁的汉密尔顿同样作出了一个史无前例的突破，他发现了力学和光学之间的类推，数学语言上说叫等式，这是一门关于光的科学。他在一封给他妹妹伊莱扎的信件中第一次提到这些想法，不过我们可以确定的是，在给他堂兄阿瑟的信件中，他确实明白地提及了这个发现的本质。

这个发现非常迷人。力学是研究物体运动的学科，比如炮弹在空气中沿着抛物线运行，钟摆从一边到另一边规律地摆动，行星围绕着太阳沿椭圆形轨道运行。而光学则是关于光信号的几何学，研究光的反射和折射，彩虹、棱镜以及望远镜片。把这两者联系在一起让人感到非常惊异，它们本质上是一样的则更让人难以置信。

不过这确是事实。这个等式直接给数学家和数学物理学家们提供了标准的研究框架，也就是所谓的汉密尔顿系统。这个系统不仅在力学和光学领域有着重要作用，其影响力还扩展到了量子理论。汉密尔

顿系统的主要特点在于从单一量推导出机械系统的运动方程，总能量被称作是系统的汉密尔顿函数。结果方程不仅包括系统各组成部分的位置，还包括它们在系统中的运动速度，也就是系统动力。另外，这个系统的美妙之处还在于，它们并不依赖于坐标的选择。美即是真，至少在数学中如此，而这里，物理学也一样美丽并正确。

*

汉密尔顿比阿贝尔和加洛瓦的幸运在于他不同寻常的天赋很早就得到了承认。所以，他在 1823 年进入爱尔兰首屈一指的都柏林大学圣三一学院就显得一点都不奇怪。在该领域的上百个申请者中，他名列第一。在三一学院，他几乎把大学设置的所有奖项收入怀中。而更加重要的是，他完成了他光学理论杰作的第一卷。

1825 年春天，汉密尔顿被爱神之箭射中，他爱上了一个名叫凯瑟琳·迪斯尼的女孩子。他愚笨地写诗来朦胧地表达爱意，可是他美丽的梦中情人却出乎意料地嫁给了一个比她大 15 岁、根本不解风情的牧师。汉密尔顿非常颓丧，尽管他是一个虔诚的宗教徒，他还是想到要把自己灌醉。不过对教徒来说这在道德上是一种罪恶。最终，他没有把自己灌得酩酊大醉，而是把自己的沮丧抒发在另一种诗作中。

汉密尔顿非常喜欢诗歌，他的朋友圈中包括一些文学界的泰斗。华兹华斯后来成了他的挚友。他还和萨缪尔·柯勒律治以及其他作家和诗人过往甚密。华兹华斯曾善意地给汉密尔顿指出他的天赋并不在诗上："你给我送来了很多诗作，这让我非常高兴……然而我们却很担心这会分散你投入在科学上的精力……我斗胆提请您考虑一下，你对诗的天性是否有利于发挥你对散文的天性……"

汉密尔顿回应道，他真正的诗作是数学，并且明智地回到科学研究之中。1827 年，汉密尔顿被全票选举为三一学院的天文学教授。因为都柏林大学三一学院的约翰·布林克雷（John Brinkley）教授辞去了天文学教授职位，他准备去爱尔兰南海岸的克罗因小镇当主教。而此时的汉密尔顿还只是一个尚未拿到学位的本科生。随后他发表了他的光学著作，这无异于给学术界投下了一枚重磅炸弹。因为这对天文学家来说是一门非常有用的科目，它是很多天文器材的制造原理。

不过此时光学与力学的联系尚处于萌芽状态。著作的主题是光线的几何学，也就是光线通过平镜或者棱镜反射后方向将会如何变化。"光线学"后来逐渐被"光波学"所取代，后者把光线也看做是一种光波，而光波拥有光线所不具备的一些性质，也就是它的衍射性质。对光波进行干扰可以使影像的边缘变得柔和，甚至可以使光线看起来沿一定的角度发生了弯曲，而这对于光线来说是不可能做到的。

在学术界，光线的几何学并不是一个新科目，早期的数学家诸如费马，或者更早的古希腊哲学家亚里士多德已经在这方面进行过思考。而现在汉密尔顿对光学作出了和勒让德对力学一样的贡献：他对几何用代数和分析进行研究，并把基于几何图形的推理代之以符号计算。

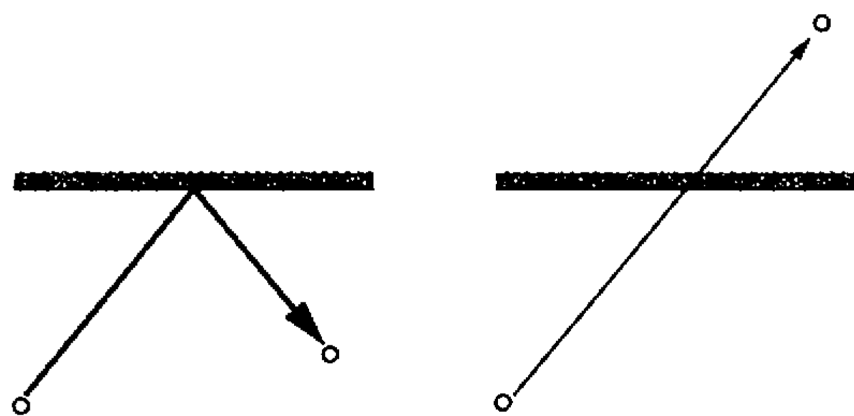
这是最重要的进步，因为这使得把不严谨的图形分析变换成为精确的数学分析成为可能。不过随后的数学家却作了艰辛的努力来改进汉密尔顿的方法，他们重新引进视像思维。不过在那时候，代数观点已经成了数学思想的一部分，并且是对视像论断的有力支持。时尚的车轮已经回到原位，不过是在一个更高的水平上，就像旋梯一样，只会越上越高。

汉密尔顿对光学的贡献是自成体系的。他沿用了大量现成结论并把它们改造成为一样的基本方法。为了代替光线学，他引进了一个单

一量，也就是系统的“特征方程”。因此，任何光学构造都可以通过这个单一方程来表示。并且，这个方程可以通过统一的方法来加以求解，因而是对光线系统和它的行为的一个完整的描述。这种方法基于一个单一的基本原理，那就是光线在穿过平镜系统、棱镜系统和透镜系统时都将遵循光线到达目的地的最短路径。

*

费马已经发现了这个原理的某些特性，并把这个原理称作最短光时原理。最简单的例子就是光线通过一个平镜来反射，下图左手边的图形表明了光线从一点出发并通过镜子的反射到达第二点的情况。早期在光学领域最伟大的发现之一就是所谓的反射定律，即光线把经过反射后有两个部分，这两个部分与镜子所成的角度相等。



最短光时原理如何导出反射定律

费马想出了一个漂亮的诀窍：把反射光线再反射到第二点上，如果这个第二点在镜子中，就如右边的图形了。这样反射的一段光线就形成了欧几里得的“等角”情形，于是在初始点和光线到达的第二点之间是一条直线。欧几里得已经证明了两点之间，直线最短。因为光线在空气中的传播是匀速的，因此，最短距离也就意味着最短时间。不反射的几何学就可以还原到左边的图形，不过论点依然成立。所

以，等角情形在逻辑上和光线在传播途中遇到一面镜子时遵循最短时间是一样的。

一个与此相关的原理，斯涅尔的折射定律，告诉我们光线从空气进入水中，或者从一种媒介进入到另一种媒介之中会发生弯曲。这似乎很简单，可以通过简单的方法加以推导，只需记住，光线在空气中的传播速度要远远高于在水中的传播速度。汉密尔顿则更进一步，他宣称，同样的最小时间原理可以应用到所有的光学系统之中。并发现这个想法存于一个单一的数学公式之中，也即特征函数。

这个数学公式极其深刻，但是在汉密尔顿的手里，这个方程却被直接应用到了实验中。汉密尔顿注意到，他的方法隐含着“圆锥折射”的存在，在这种折射中，光线在传播过程中遇到水晶体后会全部呈现出圆锥光线。1832 年，这个让当时每个从事光学研究的人都很惊讶的预测得到了汉弗莱·劳埃德（Humphry Lloyd）的验证。他使用一块水晶矿文石验证了这一预言。至此，汉密尔顿几乎在一夜之间成了科学界家喻户晓的人物。

直到 1830 年，汉密尔顿一直在考虑安定下来，并娶艾伦·德薇尔（Ellen de Vere）为妻，她对华兹华斯说，他“崇拜她的思想”。于是他开始为艾伦写诗，并正准备向她求婚，就在这时，艾伦告诉他，她永远不会离开她的家乡卡勒平原（Curragh）。汉密尔顿把这理解为她是在委婉地拒绝，事实证明，他的这个理解是正确的，因为仅仅过了一年，艾伦就嫁给了别人并搬到了其他地方。

最终，他和居住在气象台附近的当地少女海伦结婚了。汉密尔顿把她描述为一个“完全不出众的”女孩。蜜月就如灾难一样，汉密尔顿潜心光学研究，海伦则一病不起。1834 年，他们的儿子威廉·埃德文出生，随后的一年中，海伦大部分时间都处在出走状态，1835 年，

他们的第二个儿子阿尔博奇德·亨利出生，不过汉密尔顿和海伦的婚姻也基本上走到了尽头。

*

后世人认为汉密尔顿的机械光学类推是他最伟大的发现。但是对汉密尔顿来说，直到去世，他都处在与日俱增的迷惑之中，属于他的伟大荣耀却是和光学完全不同的另外一个领域：四元数。

四元数是一种代数结构，和复数很相似。汉密尔顿坚信，四元数是解决物理学深层问题的钥匙。事实上，在他生命的最后几年，他认为四元数是打开一切的钥匙。然而历史似乎并不同意这种看法，在随后的一个世纪中，四元数逐渐淡出了人们的视线，仅仅成为抽象代数的一个模糊的源头，几乎没有什么重要的应用。

直到最近，四元数才迎来了复兴。尽管它们远没有得到汉密尔顿所希望得到的重视，不过四元数正逐渐被看做是一种重要的数学结构的源头。事实证明，四元数是很特别的怪物，在现代理论物理中必不可少。

四元数刚被发现的时候，在代数领域引起了一场巨大变革。四元数打破了一种重要的代数规则。在约 20 年中，几乎所有的代数规则都以同样的方式被打破，既导出了很多有益的新结论，也出现了很多难以突破的死路。19 世纪中期的数学家们所认为不可侵犯的数学规则不过是一些使得代数学家的分析变得简单的假设而已，它们并不能够满足数学本身的更高要求。

在勇敢的新“后加洛瓦”世界里，代数已经不仅仅是代替方程中的数字的符号，而是方程的深层结构——不是数字而是过程、变换与

对称。这些大胆的创新改变了数学的面貌，使得数学更加抽象，更一般且更有力量。整个领域具有一种奇怪的，无法言喻的美丽。

直到文艺复兴时期，博洛尼亚的数学家们才开始探讨 −1 是否存在有意义的平方根，是否数学中出现的所有数字都同属于一个单一的系统。今天，作为历史混淆数学与现实的后果，这个系统被叫做“实数”。然而这个名字是不幸的，因为这意味着，这些数字并非是人们对宇宙的理解，而是对宇宙的虚构。它们并不“真实”。它们并不比其他在过去的 150 年里通过人们的想像力构建出来的其他数字系统真实；不过它们确实比任何新创造出来的数字系统与现实联系得都更加直接。它们与一个理想的度量方式紧密相关。

实数的实质就是十进制。这并不是从它作为方便计算的记号来说的；实际上，十进制的实数反映的是十进制本身的内在性质。实数由简单和清楚的源头演化而来。首先，人类把它勉勉强强地限制在自然数系统中，比如 0，1，2，3，4，等等。我之所以这么说，是因为在早期，这些数字中的很多根本不成其为数字。曾经有一段时间，古希腊人拒绝承认 2 是一个数字，因为它相对于数量来说太小了。因此，数字从 3 开始。最终他们承认 2 和 3，4，5 是一样的数字，但是却不承认 1。毕竟，如果一个人说他有“数头牛”，可结果他只有 1 头牛的话，这会被认为是撒谎而被加以责罚。“数头”自然意味着多于 1，因此排除了只有一头。

但是随着记号系统的发展，越来越明显的是数字 1 在计算过程中的作用就和其他大的数字一样。所以，1 变成了一个数字，不过却是一个很特别，也很小的数字。在某些方面来说，1 反而是一个最重要的数字，因为它是数字的开始。把许多 1 加在一起，你可以得到任何其他数字。事实上人们也确实是这样做的，在某段时间内，7 通常是

被写成 7 根竖线。

很久以后，印度的数学家承认在数字 1 之前还存在一个更加重要的数字。1 并不是数字的起点，他们从零开始，写作 0。又过了很久，人们才把小于 0 的负数加进系统，事实证明这很有用处。负的数字加进来以后，人们创造了整数的概念：……−3，−2，−1，0，1，2，3，……不过数字的发展并未停止。

所有数字的问题是，它们仍然难以描述许多有用的性质。举例来说，一个农民卖掉粮食，可能遇到粮食在 1 袋和 2 袋之间的情形。如果正好是一半，那么就是 $1^1/_2$，或者可能稍微少点，比如 $1^1/_3$，或者多点，$1^2/_3$。因此，分数被创造出来了，不过记法却很多。分数插进了整个数字系统之中。极其复杂的分数在数字系统中的插入非常妙，就如我们已经在古巴比伦算术中见过的那样。任何数字都可以通过分数来表示。

这时出现了毕达哥拉斯和以他名字命名的定理。这个定理的一个重要推论就是单位正方形的对角线长度的平方是 2。那就是说，对角线的长度是 $\sqrt{2}$。这样的一个数字一定要存在，因为你一定可以画出一个单位正方形，并且这个正方形的对角线一定是存在的，那么对角线就一定会有长度。就像希波索斯的悲剧结局一样，不管怎么计算，$\sqrt{2}$ 都不是一个分数。它是无理数。在分数系统中，还有很多看不到的空白部分需要新的数字来填满。

*

最终，这个过程看起来是停止了。希腊人随后抛弃了数字主题而钟情于几何。但是在 1585 年，一个住在布鲁日的名叫西蒙·斯蒂

文（Simon Stevin）的弗莱芒数学家兼工程师被威廉一世（William the Silent）聘请为他儿子莫里斯（Maurice of Nassau）的家庭教师。斯蒂文成了堤防的巡视员，部队的军需官和财政部长。这些任命，尤其是最后的两项促使他去寻找一个恰当的记录方式，于是他引入了意大利人的书记方式。他在寻找一个表述分数的方法，这个方法要同时具有印度—阿拉伯记号的灵活性，还要具有巴比伦六十进位的精确性。最终，斯蒂文得到了基于十进位的系统，即十进制，这和巴比伦六十进制系统比较相似。

斯蒂文发表了一篇文章来描述其新的记数系统。他非常用心地对其进行推销，甚至写道："这是一个由从事实际工作的人开创的彻底胜利，从事实际工作的人发现这种计数系统非常有用，以至于他们不得不不情愿地抛弃他们以往的习惯而采用这个新的系统。"并且，他宣称，他的十进制系统"可以满足一切在商业中遇到的计算，而且只需整数计算不需要任何分数的帮助"。

斯蒂文的记号并没有使用今天十进制意义上的记法，不过二者却是直接相关的。我们写 3.1416，在斯蒂文那里写作 3⓪1①4②1③6④。记号⓪代表整数，①代表是十分之一，②是一百分之一，等等。当人们逐渐熟悉了这种记法之后，他们就把①②等省略，仅仅留下⓪，变成十进位的那个点。

在十进位系统中，我们并不能写出 $\sqrt{2}$ 是多少，因为一旦开始我们就无法停下来。同样，在十进位中我们也不能写出 $^1/_3$ 是多少，它接近 0.33，不过 0.333 更接近一点，0.3333 当然更加接近。要想有精确一点的表述，只能允许这个 3 的篇幅无限长。但如果可以接受的话，我们也可以比较精确地写下 $\sqrt{2}$ 是多少。

如果可以接受无穷小数，那么实数系统就变得完整了。可以根据

商人或者数学家的需要写出任何精确的数字来。任何一个构造出来的记数方式，都可以以十进制的方式进行表示。如果需要写下一个负数，十进制的方式也可以轻易做到。因为不再可能需要其他种类的数字，因而也不存在空白需要去填补。

*

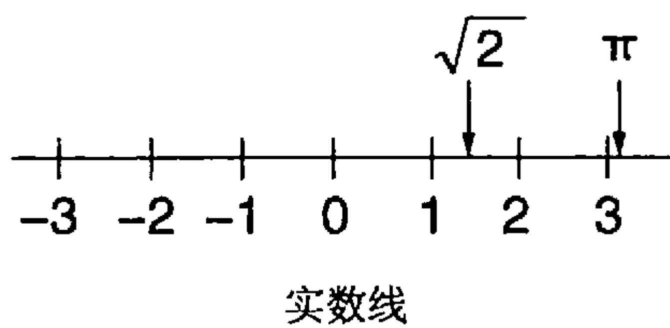
只有卡尔达诺的三次方程公式看起来似乎可以告诉我们一些事情，不过却是极其让人迷惑的。如果你从一个看起来并不麻烦的三次方程出发，假设你知道这个方程一定有根，可是卡尔达诺的公式却并不会给出解析解，相反，使用其公式的话反而会出现要开三次方的更复杂的表示形式，甚至更加复杂的事情是，需要求负数的平方根，这是根本不可能做到的。毕达哥拉斯学派受阻于 2 的平方根，而 −1 的平方根则更加令人沮丧。

几百年来，给 −1 的平方根以意义一直挑战着数学家们的智力。如果你想得到一个数，这个数的平方是 −1，也就是说 −1 的平方根，这个数一定是虚数，所以根本不存在这样一个解。除了笛卡尔之外，还有很多人想到了这点。1637 年，笛卡尔从实数中找出了“虚数”的存在，并且他坚信，正是虚数的存在使得解不存在。牛顿几乎也说明了同样的事情。但是所有这些渊博之士都是在了解邦贝利的贡献之前作出的猜想。邦贝利在几个世纪之前就注意到，虚数的存在可能是方程存在解的一个信号。不过这个信号很难理解。

1673 年，出生于阿什福德的英国数学家约翰·沃利斯（John Wallis）作出了惊人的突破。他的出生地阿什福德与我的家乡坎特不过十五英里。他发现了一种简单方法，以坐标系中的点来表示虚数，

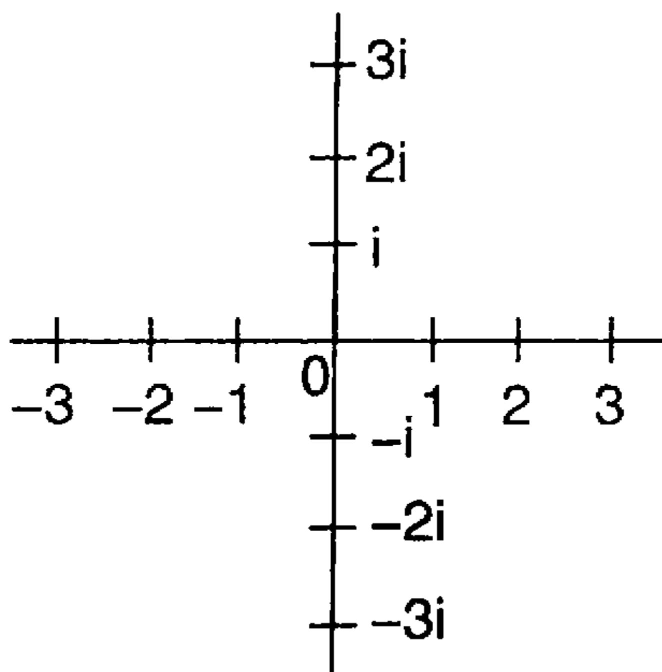
甚至可以表示出那些较为复杂的具有实部和虚部的数字。第一步是使用现在较为熟悉的实数线的概念，这条线向两边无限延伸，0 处于实数线的中央，正实数向右依次排开，负实数向左排开。

每个实数都可以在这条线对应的位置上。每一个后续的十进位小数都是把单位长度分成十份、一百份、一千份等相等的部分，这都不存在障碍。像诸如 $\sqrt{2}$ 这样的数字，可以将其精确到我们需要的程度，肯定是在 1 和 2 之间，较为接近 1.5。数字 π 的位置是在 3 的右边一点，依次类推。



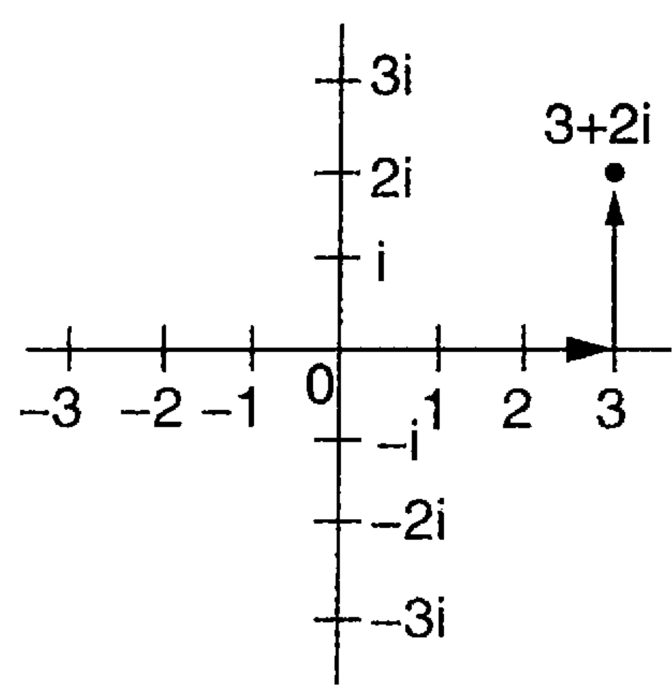
实数线

但是问题是 $\sqrt{-1}$ 在哪里呢？在实数线上没有它的位置，它既不是正数，也是不是负数，因此，不能在 0 的左右对其进行表示。所以，沃利斯将其放在了其他地方进行表示。他引入了第二条实数线来包括虚部，也就是一系列的 i，并将其放在实数线的右边。用出面点的话来说，这种方法是一种“水平思考”。



两条成直角的实数线

这两条数字线，实数线和虚数线要相交于 0 点。可以较为容易地证明，如果所有的数字都有意义，那么 0 乘以 i 一定等于 0，所以实数线和虚数线正好相交于 0 点。一个复数有两个部分组成，即实部和虚部。为把这个数字放在坐标里面，沃利斯让读者把实部在水平的实数线上表示出来，而把虚部用垂线表示出来，和虚数线平行。



沃利斯的复数平面

这个办法完全解决了赋予虚数和复数以意义的问题。很简单但却具有决定性，绝对是神来之笔。但在当时，这种方法却被完全忽视了。

*

尽管没有得到公众承认，沃利斯的突破却渗入了数学家的意识。因为数学家们下意识地使用了与沃利斯有着直接关系的基本想法，那就是不存在复数线，但是却存在复数坐标。

随着数学的用途变得越来越广，数学家们开始试着进行更复杂的运算。1702 年，约翰 · 伯努利（Johann Bernoulli）试着解决一个微积

分问题，他需要求出一个复数的对数值。直到1712年，伯努利和莱布尼兹才解决了一个核心的问题：负数的对数是什么？如果你能够解决这个问题，那么你就能够解决任何一个复数的对数形式。因为，一个数字的平方根的对数正好是那个数字对数的一半。所以，i 的对数实际上是 -1 的对数的一半。但 -1 的对数是多少呢？

争论的关键问题很简单。莱布尼兹确信 -1 的对数是复数，而伯努利则认为是实数。伯努利的论点基于简单的微积分之上，而莱布尼兹则认为无论是从结论还是方法上，这都讲不通。1749年，欧拉解决了这个争论，不过结果却很倾向于莱布尼兹。他指出，伯努利的方法忽略了某些东西，他的微积分计算中包含着一个“任意常数”，而在他的复数微积分的热情中，他假定这个任意常数是0。实际上并非如此，这个常数是一个虚数。这个小小的疏忽解释了伯努利的结论和莱布尼茨的结论之间不同的原因。

数学的复数化步伐因此得到了极大加快。许多始于实数领域的研究被拓展到复数领域。1797年，一个名叫卡斯帕·维塞尔（Caspar Wessel）的挪威人发表了一种用坐标系中的点表示复数的方法。卡斯帕出生于一个教会的牧师家庭，是家里14个孩子中的老六。在当时，挪威被丹麦统治着，并且没有大学，所以卡斯帕在1761年去了哥本哈根大学。他和他的哥哥欧里学习法律，为分担家庭经济压力，欧里随后做了测量员，不久卡斯帕成了他的助手。

在做测量员期间，卡斯帕发明了一种坐标的几何表示方法，尤其是对坐标系的线和方向用复数形式进行表示的方法。反过来说，他的思想可以看做是在坐标几何中表示复数的方法。在1797年，他把他的这一成果上报给了丹麦皇家科学院，这是卡斯帕一生中惟一的一篇关于数学的研究文献。

但是几乎没有顶尖的数学家可以阅读丹麦语，因此直到一个世纪之后，这项成果被翻译成法语才为人所知。而在同时，法国数学家让－罗贝尔·阿根（Jean-Robert Argand）也独立地发明了这种想法并于1806年发表了其研究成果。1811年，高斯也独立发现，复数可以作为坐标系中的一个点。数学术语诸如“阿根图”、“维塞尔坐标”，以及“高斯坐标”开始流行，不同国家的人会使用各自偏爱的叫法。

决定性的一步则是由汉密尔顿来迈出。1837年，大约在卡尔达诺方程认为虚数可能有用的大约三百年之后，汉密尔顿把复数中的几何元素精简为纯粹的代数。他的思想很简单，同样的思想也隐含在沃利斯的建议中，并且与维塞尔、阿根以及高斯的想法相同，只是他们都没有将其明确化。

汉密尔顿认为，从几何意义上说，坐标系中的一个点可以看做是一个实数对，坐标表示为（x, y）。如果观察沃利斯图形（或者维塞尔、阿根，或者高斯的图形），会发现 x 是一个数的实部，而 y 是这个数的虚部。一个复数 $x+yi$ 仅仅是一个实数对（x, y）而已。你甚至可以应用法则来对这些数字进行相加和相乘，而主要的步骤是要注意，i 对应着数对（0, 1），那么（0, 1）×（0, 1）一定等于（−1, 0）。对此，高斯在1831年给匈牙利几何学家沃尔夫冈·波尔约的一封信中精确地揭示了同样的想法。狐狸又一次成功地掩盖了他的伎俩，因此完全没有被发现。

问题最终得到了解决。一个复数不过是一个实数对而已，并通过应用简单的规则来进行变换。因为，一个实数对与单个实数没有什么区别，因此实数和复数都和现实紧密联系，所谓的虚数实际上起了误导作用。

今天的观点则是相当不同的：“实数”是一个有误导性的概念。

无论是实数还是虚数都不过是人类臆造的事物而已。

*

对汉密尔顿解决的三百年来难题的反应显然被减弱了。曾经，数学家们把复数的概念应用到强有力的一致性的理论中，并且非常担心复数的存在会变得不重要。但汉密尔顿的数对则无论如何都具有重要意义。即使复数如今已经不是让感到我们兴奋的源泉，但从旧数字系统之上建立新的数字系统的思维仍然存于数学家的意识深处。

复数最终被证明不仅在代数和基本微积分中有用，它们也是解决流体流动、热力、重力、声音，乃至数学物理学所有领域的有力工具。但是，复数也有一个主要的局限，那就是只能在二维空间中解决这些问题，而不是我们生活于其中的三维空间。不过，有些问题，诸如鼓皮的运动以及流体的流动都可以简化为二维空间，因此，并不都是坏消息。不过数学家们不能将其使用的复数方法从二维空间拓展到三维空间，这一点正在使他们变得沮丧和恼怒。

是否存在一种尚未被发现的把二维数字系统拓展为三维数字系统的方法？汉密尔顿把复数标准化为一个实数对的方法已经比较接近于这样一种考虑：那就是试图建立一种三重数字系统（x, y, z）。不过问题是，没有人曾经建立过三维代数。汉密尔顿决定一试。

加进三重性并不复杂，可以从复数中得到线索，只需要增加相应的坐标就可以了。这种算法今天被称作“增加向量”，它遵循着良好的规则，并且只有一种方法来进行。

困难之处则在于乘法。即使对复数，乘法也不像加法。你不能把两对实数通过分别将它们的第一部分和第二部分相乘来进行。如果你

真的这样做了的话，会有很多便利，但是两个致命的不便也会出现。

第一个是，不会再有 -1 的平方根。

第二个是你可以通过把两个非 0 的数字相乘而得到 0。这些 0 的除数可以和普通的代数方法一样应用，比如解方程。

对于复数来说，我们可以通过选择不那么明显的规则来进行相乘，从而解除这个障碍，那就是汉密尔顿使用的方法。但是，当他尝试将相似的技巧应用在三维空间上时，却得到了让人震惊的结果——难以避免一些致命的缺陷。他得到了 -1 的平方根，但是仅仅是通过引入 0 的除数的时候才会出现。而无论他怎么做，想把 0 的除数去掉都是完全不可实现的。

*

如果你觉得这听起来似乎是有点尝试解五次方程的意思，那你就想错了。如果许多优秀的数学家尝试去做某事，然而失败了，那么很有可能是因为那件事根本就是不可能实现的。如果说数学家真的教会我们什么的话，那就是许多问题根本就是没有解的。你无法找到一个平方是 2 的分数；你无法用圆规和直尺来三等分一个几何角；你无法从根本上求解一个五次方程。数学是有自己的极限的。也许你无法构建一个拥有你想要的所有良好性质的三维代数。

如果你想要知道是否真是如此，那么这正好是一个好的研究命题。首先，你应该知道你的三维代数应该具有什么性质，然后你要分析这些性质的逻辑后果。假定从这个试验中获得了足够的信息，那么你就能找出这种代数所应该具有但实际上又没有的性质，并且推断其并不存在的原因。

至少这是你今天应该尝试做的事情。汉密尔顿的方法并不是那么系统。他假设他的代数应该具备所有良好的性质，然后又突然发现某个性质是应该免除的。更加重要的是，他认识到，三维代数并不可能。他发现，最可能得到的是四维而不是三维。

回到那些难懂的代数规则上来。当数学家们进行代数运算的时候，他们把代数符号以系统化的方式进行重新排列。回忆一下 algebra 的阿拉伯名字，“al-jabr”原意是“重建”，而我们今天将其叫做“把某项移到方程的另一边并且改变其符号”。仅仅是在最近 150 年里，数学家们才费力地将许多操作规则明确地写了出来，并从中推出了其他一些著名的逻辑结论。这套对代数进行公理化的方法就像欧几里得在几何领域做的工作一样，而得到这种思想却花费了数学家们两千年的时间。

为做好准备，我们可以集中于这些规则中的三条，这三条都与相乘有关。（相加的规则很相似不过更加直观；相乘是很多事情开始走向梨型 [pear-shaped] 的起点。）孩子们在学习乘法口诀的时候最终会发现很多计算只需反过来就可以了。3 乘以 4 等于 12，4 乘以 3 也等于 12。如果你把两个数字相乘，无论哪个数字在前面，最终都会得到相同的结果。原因在于所谓的交换律，用符号来表示就是对任何两个数字 a 和 b，都有 $ab=ba$。这个法则在复数系统中也是成立的。你可以通过检验汉密尔顿方程来对此进行证明。

更加精妙的法则是“结合律”，当你把三个数字顺序相乘的时候，从哪里开始相乘并没有不同。举例来说，假设我想计算 $2\times3\times5$，我可以先把 2 和 3 相乘，得到 6，然后再用 6 乘以 5。另外的方法是，我可以先把 3 和 5 相乘，得到 15，然后再用 2 乘以 15。无论哪种方法都可以得到同样的结果，也就是 30。结合律决定了它们会得到相同

的结果。用符号来表示就是 $(ab)c=a(bc)$，括号表明了相乘的两种方法。这个定律也可以同时应用到实数和复数，也可以通过汉密尔顿方程进行证明。

最后一个非常有用的法则是分配律。虽然你可以发现在教科书中，这条法则被表述为存在一个可乘的倒数，也就是说，你总是可以同时在方程两边除以一个非 0 的数。要求是非 0 数的理由很充分，其中之一就是除以 0 是没有意义的。

我们前面已经看到，你可以通过使用一个明显的乘式来构造一个三维代数，这个系统满足交换律和结合律，但是却不满足分配律。

在进行了几个小时的计算和搜寻之后，汉密尔顿的巨大灵感突然闪现，那就是有可能构造一个新的数字系统，在这个系统之中，结合律和分配律都可以进行应用，但是却不能使用交换律。即使如此，仍不能进行三个实数的计算，你必须使用四个实数。实际上并不存在有意义的三维代数，但却存在相对较好的四维代数。这是这个种类的惟一类型，只有一个方面没有达到理想系统的要求，那就是交换律不能够成立。

这有什么关系吗？汉密尔顿思考中最大的心理障碍就是他认为交换律是很关键的。1843 年 10 月 16 日，所有的一切几乎在瞬间发生改变，他突然明白了如何去进行四维数的相乘。而在当时，他正和妻子沿着皇家运河的大路去位于都柏林的爱尔兰皇家科学院开会。他的潜意识正在被三维代数的问题所搅扰，但灵感在瞬间出现。“当时，我的思维就像突然通路了一样”，他在随后的信件中表示，“这个灵感的火花就是我一直使用的关于 i, j, k 的基本方程。”

他立刻把他的方程刻在了布罗汉姆桥（Broome Bridge，他把该桥称作是“Brougham”）上。这座桥留了下来，虽然有一个纪念性的

标志，但是当时的镌刻痕迹却没有得到保存。不过方程式却流传了下来：

$$i^2=j^2=k^2=ijk=-1$$

这是些非常美丽的方程式，并且具有对称性。但是你很可能会问，四维数在哪里呢？

复数可以写成数对（x, y），但是它们通常写成 $x+yi$，这里 $i=\sqrt{-1}$。同样，汉密尔顿头脑中的数字可以被写成四维数（x, y, z, w），或者写成 $x+yi+jz+kw$。汉密尔顿公式使用的是第二个表达法，如果你具有变通的思维的话，你也许会更加倾向于使用四维数。

汉密尔顿把他的新数字叫做四元数。他证明了这些数字满足结合律，并且更加重要的是——这也是后来才发现的——它们也满足分配律，但是不满足交换律。四元数的相乘法则意味着 $ij=k$，但是 $ji=-k$。

四元数系统包含着一组复数，也就是具有 $x+yi$ 形式的四元数。汉密尔顿公式表明，-1 并不只有两个平方根，i 和 $-i$，也有 j 和 $-j$，k 和 $-k$。实际上，在四元数系统中，有无穷多个 -1 的平方根。

顺着交换律的思路，我们也发现，四元方程的根不是两个。幸运的是，直到我们发明四元数的时候，代数的关注点已经从求解方程的根上转移到其他方面去了。四元数的好处远远超过了其缺陷，你需要的仅仅是习惯它们而已。

*

1845 年，托马斯 · 迪斯尼（Thomas Disney）带着他的女儿去拜访汉密尔顿，威廉在童年的时候很爱凯瑟琳（Catherine）。那个时候，她失去了她的第一任丈夫并且已经再婚。这一相逢使得汉密尔顿再度

陷入对往日的痛苦回忆之中，他更严重地酗酒以使自己摆脱悲伤。在都柏林的一次科学晚宴上，他喝得酩酊大醉并且丑态百出。后来他坐马车回家后，决定两年之内滴酒不沾。但当天文学家乔治·阿莱伊（George Ariy）奚落他戒酒时，汉密尔顿开始以更加大量的喝酒对他进行回应。从那以后，他就成了一个积重难返的酒鬼了。

两个叔叔相继去世，一个同事同时也是他的好朋友自杀，加上凯瑟琳再次给他写信，这些都让汉密尔顿感到非常沮丧和难过。她很快认识到，她的这种行为并不是一个正直的已婚女人应该做的事情，并且一度很想自杀。于是她暂时地离开了她的丈夫，回去同母亲住在一起。

汉密尔顿则通过凯瑟琳的亲戚不断地寄信给她。直到1853年，她更新了她的联系方式并且送给了他一个小礼物。汉密尔顿则亲自去看望她，并且送给她一本他写作的关于四元数的书稿。两个星期之后，她离开了人世。这给汉密尔顿带来了巨大的悲痛，从此他的生活更加无序和混乱。1865年，他因酗酒过度得了痛风而离开了人世，在他的数学研究书稿之中，还有尚未吃掉的食物。

*

汉密尔顿相信，四元数是代数和物理学的圣杯，真正地把复数扩展到了高维空间，并且是空间几何和物理学的关键。当然，空间是三维的，而四元数则是四维的，但是汉密尔顿却发现了一个自然的三维子系统。这些就是“虚构的”四元数 $bi+cj+dk$。从几何上说，这些符号 i，j，k 可以理解为空间中沿着三个相互垂直的坐标进行的旋转，但是却有一些微妙之处，那就是你必须在一个周角是720度的几何中

进行研究，而不是周角为360度的几何之中。除了这点之外，你就可以知道汉密尔顿为什么认为四元数对空间几何和物理学如此重要了。

这失掉的“真实的”四元数就像实数一样。你不可能把他们都清除掉，因为当你进行代数运算时，它们可能会突然出现，即使是仅从“虚构的”四元数开始运算。如果可能独立地出现在虚四元数中，那么也就很可能存在有意义的三维代数，这样汉密尔顿的探讨就成功了。四元数的四维系统是另一个美妙之处，自然的三维系统规则地嵌入其中，并且就像纯粹的三维代数那样有用。

汉密尔顿把他的余生都献给了四元数，他发展了四元数的数学形式并推动其应用到物理学中。但是只有很少的学者对此加以赞扬。他们建立了一个关于四元数的学派，在汉密尔顿去世之后，这个学派由爱丁堡的数学家皮特·泰特（Peter Tait）以及哈佛大学的数学家本杰明·皮尔斯（Benjamin Peirce）来领军。

然而其他学者并不喜欢四元数，一方面是因为四元数的随意性，更加重要的是因为他们认为他们找到了更好的替代品。在这些不同意见者之中，最著名的是普鲁士数学家赫尔曼·格拉斯曼（Hermann Grassmann）以及美国物理化学家吉布斯（Josiah Willard Gibbs），他们今天被公认为是向量代数的创始人，他们俩都发明了有用的高维代数。在他们的工作中，并不存在四维的极限或者虚四元数的三维极限。这些向量系统的代数性质并不如汉密尔顿的四元数雅致。比如，你不能够进行两个向量的相除。但是格拉斯曼和吉布斯更喜欢这种一般性的概念，即使这些概念可能缺少某些良好性质。确实不能进行两个向量的相除，但是他们也不关心把两个向量相除会得到什么。

汉密尔顿至死都认为四元数是他对数学和科学作出的最重要的贡献。但是在接下来的几百年间，除了泰特和皮尔斯之外，几乎没有人

同意这种看法。在维多利亚时代的代数学中，四元数成了一潭静止不动的死水。如果你想找到例子来证明那个时代数学的思想贫乏状态，四元数的遭遇就是一个恰当的反映。即使在大学的理论数学课程里，四元数也从来没有出现过，或者至多只是作为猜想而存在。贝尔（E. T. Bell）写道：“汉密尔顿最惨痛的悲剧不是酒精或者婚姻，而是他固执地认为他的四元数是数学和物理世界的根本。历史已经表明，当汉密尔顿坚信他的四元数的意义就像19世纪中期的微分学终结了17世纪一样重要时，他实际上悲剧性地欺骗了他自己。还从来没有一个伟大的数学家错得如此悲哀。”但真是这样吗？

四元数虽然没有沿着汉密尔顿所说的方向发展，但是其重要性却在科学界逐年增强。它们已经成为数学的绝对基础，我们也将看到，四元数及其推广将成为物理学的基础。汉密尔顿的发现为现代代数和现代数学物理学的快速发展打开了大门。

真实的情况是，从来没有一个半吊子的历史学家错得如此悲哀。

*

汉密尔顿也许是过于夸大了四元数的应用，并且有时候把四元数不适宜地应用进某些技巧之中，但是历史越来越证明，他对四元数重要作用的信心却并不过分。在很多看起来几乎不可能的地方，四元数都会突然出现。它们可以通过一些有道理并相对简单的运算法则进行划分——当然，除了一条重要的法则不能运用——由此，四元数构成了一类具有那些法则的特殊的数学系统。

不过应该对此表述进行细致分析。

对所有人来说都很熟悉的数字系统是实数系统。你可以对实数进

行加减乘除，并且总是会得到另外的实数。当然，0不能做除数，除此之外，你可以把这些运算法则随意应用，并且肯定不会脱离实数系统。

数学家把这样的系统叫做一个"域"。当然数学中也有很多其他的域，比如有理数和复数。但是实数域却很特别，它是惟一具有有序性和完备性的域。

"有序的"意味着数字是以线性顺序排列的，实数是沿着一条线的顺序排开的，负数在左，正数在右。数学中也存在着其他的有序域，比如有理数系统。但是与其他的有序域不同的是，实数域又是完备的。这个额外的特点保证了$\sqrt{2}$和π等数字的存在。本质上说，完备性使得无穷小数具有意义。

可以证明，实数是惟一的完备有序域。这也是实数在数学中占据核心地位的原因。实数域也是惟一使得某些运算，比如"大于"（Greater than）和基本的微积分操作得以执行的域。

复数通过增加某些新的类型的数字拓展了实数，比如 -1 的平方根。但是我们为求解 -1 的平方根而付出的代价则是丧失了有序性。复数是一个完备的系统，但确是沿着坐标系而不是沿着一个单调的有序数列而展开的。

坐标系是一个二维系统，而2是一个有穷整数。复数是惟一一个含有实数和有穷维数的域，而不像实数域那样是一维的。这意味着复数是惟一的。对很多重要的目的而言，复数是为一个能够满足这种要求的新发明。它们的惟一性保证了它们的不可或缺性。

当我们试图去拓展复数系统的时候，去增加维数并保证很多重要的运算法则继续可以应用的时候，四元数就出现了。而那些我们试图去保持的运算法则也就是通常的加减法，乘法的大部分性质，以及

被非零数去除的可能性。不过代价却更加高昂了，这也是让汉密尔顿心痛的原因，你必须抛弃乘法的交换律。你只能把这当做是成规来接受，然后继续下去。不过当你真正习惯了以后，你可能就会反问，为什么乘法的交换律一定要在任何条件下都成立呢？并且你肯定会认为，在复数系统中，乘法的交换律成立简直就是一个小奇迹。

任何具有这些交织在一起的性质的系统，不论交换律成立与否，都被叫做是可除代数。

实数和复数都是可除代数，因为我们并没有排除乘法的交换性，我们只是没有要求乘法的交换性而已。任意域都是可除代数。但是一些可除代数却不一定是域，第一个被人们发现的就是四元数。1898年，胡尔维兹（Adolf Hurwitz）证明了，四元数系统是惟一的。四元数是惟一的具有有穷维数的可除代数，是包含实数但是又不等同于实数或者复数的系统。

这就出现了一个让人惊奇的类型。实数、复数，以及四元数的维度分别是1，2，和4。这看起来似乎是某个序列的开始，也就是2的幂的序列，一个自然的延伸应该是8，16，32，等等。

难道真的存在某种具有这样的维数的代数系统吗？

答案是存在也不存在。要过一会儿你才能知道原因，因为关于对称的故事现在进入了一个新的阶段：与微分方程联系在一起的阶段。微分方程是对物理世界进行建模的最广为使用的方式，也是许多出色的物理学家表述他们对自然法则认识的基本语言。

理论的最深刻部分再一次被归结为对称，不过却是通过新的方式实现的。现在，对称群不再是有穷的，而是“连续的”。通过这种前所未有且极具影响力的研究方法，数学将得到极大的丰富。

冒牌士兵和虚弱的书虫

M. S. 李（Marius Sophus Lie）走上科学之路完全是因为他的视力不好而无法胜任任何军职。1865 年，当他被要求从克里斯蒂安尼亚大学毕业的时候，他实际上只上了很有限的几门数学课，这其中包括一门由挪威人卢德维格·西罗（Ludwig Sylow）讲授的加洛瓦理论，但是对这门学科他并未表现出任何过人的天赋。他曾一度非常迷茫，他知道自己想要一个以学术工作为主的职业，但是却不确定是应该在植物学领域选择，还是动物学领域，或者是天文学领域。

不过图书馆的借书记录却显示，他之后越来越多地借一些数学书籍。1867 年的一个午夜，他突然间明白了自己的一生应该从事什么样的工作。于是他激动地给他的朋友恩斯特·莫兹菲特（Ernst Motzfeldt）打电话，莫兹菲特从梦中惊醒，李在电话那头激动地喊道，“我终于找到它了，它是如此简单啊。”

他所找到的是一种思考几何的新方法。

于是，李开始研究历史上所有伟大的几何学家的著作，诸如德

国人普吕克（Julius Plücker），以及法国数学家蓬斯莱（Jean-victor Poncelet）。从普吕克那里，他体会到几何的关键元素并非欧几里得的点，而是诸如线、坐标系以及圆周之类的其他东西。1869年，他自费发表了一篇文章来阐明自己的思想。像他的前辈数学家加洛瓦和阿贝尔一样，他的思想对于守旧者来说过于前卫和具有革命性，以至于一些中规中矩的杂志均不愿意发表他的研究成果。但是恩斯特不想让他的朋友感到沮丧，并鼓励他坚持几何学的研究。最后，李的一篇文章终于在一个很出名的杂志上得到了发表，并为学术界接受。这也为他赢得了一项奖学金。于是李开始有钱进行旅行，拜访顶尖的数学家并和他们讨论他的思想。他来到产生很多普鲁士和德国数学家的福地哥廷根和柏林。在那里他拜访了代数学家克罗内克（Leopold Kroneker）和库默尔（Ernest Kummer），以及数学分析学家魏尔斯特拉斯（Karl Weierstrass）。在拜访过程中，他被库迈尔的数学研究方法所打动，而不是魏尔斯特拉斯的研究方法。

而最重要的会面则是他在柏林见到了克雷恩，恰巧克雷恩正是普吕克的学生，而普吕克又是李最想效仿的数学家。李和克雷恩具有非常相似的数学背景，但是对数学的嗜好却相去甚远。克雷恩本质上是一个使用几何工具的代数学家，喜欢钻研一些具有内在美的特殊问题。而李本质上是一个数学分析学家，对具有广泛应用性的一般理论着迷。但具有讽刺意味的是，正是李的一般性理论提出了很多数学理论上的特殊结构，而这些特殊结构今天仍然极具美感，并且异乎寻常地深奥，而且大部分用代数形式表示。如果不是李在一般理论上的推进，这些特殊问题可能至今都不会得到解决。如果你试着去理解一门数学学科的所有问题，并且能够成功做到的话，你很可能会发现，许多数学问题具有很不寻常的特点。

1870 年，李和克雷恩在巴黎再次相见，在那儿，若当把李的研究兴趣转到了群理论的原因上。越来越多的研究已经发现，几何理论和群理论实际上是一枚硬币的两面，但是，这个思想的全面形成却经历了漫长的时间。李和克雷恩一起做了一些工作，他们试着去把几何和理论联系起来。最终，克雷恩在他 1872 年的著作“爱尔兰根纲领”（Erlangen Program）中明确提出，几何和群论本质上是一样的。

用现在的语言来说，这个思想听起来非常简单，并且应该早就被发现。与任意一种几何相适应的群在几何上都是对称群。反过来说就是，与任何群相适应的几何都使群成对称。因此，几何被群中那些不变的东西限定着。

举例来说，欧几里得几何中的对称就是那些保持长度、角度、线和圆周的平面变换。这是在平面上保持严格运动的群。换句话说，任何在保持严格运动的过程中不发生变化的东西都将自然地落在欧几里得几何之内。非欧几里得几何仅仅是使用了不同的变换群而已。

那么为什么要费劲地把几何变换为群理论呢？这是因为这样你可以有两种方法来思考几何问题，并且有两种方法来思考群的问题。有的时候，一种方法可能较好理解，但某些时候可能另一种方法更好理解。对任何事情来说，有两种看法总比只有一种看法好。

*

法国和普鲁士之间的关系急剧恶化，拿破仑三世认为他可以通过向普鲁士宣战来提升他的民意支持度。俾斯麦向法国发出了一份措辞强硬的电报，1870 年 7 月 19 日，普法战争正式爆发。克雷恩当时正在巴黎，作为一个普鲁士人，他出于谨慎返回了柏林。

然而李是一名挪威人，停留在巴黎对他来说没有任何危险，他充分地享受着自己的旅行。但是，当他意识到法国可能战败，而德国军队正在向梅斯挺进的时候，他改变了主意。尽管是一个中立国的国民，在潜在的战争区停留仍然是一件极度危险的事情。

李决定继续进行徒步旅行，开始向着意大利进军。然而还没有走多远，法国当局就在距巴黎南部25公里的枫丹白露抓住了他，他随身带着很多写满难以理解的符号的文件。因为这些符号，他被当做德国的间谍投入监狱。在当时法国著名的数学家达布（Gaston Darboux）的干涉下，法国当局才最终认定那些符号是数学。李得以释放，而法国军队也在这个时候投降了，德国人封锁了巴黎。李再次决定去意大利，这次他成功了。他从意大利返回了挪威，途中，他顺便拜访了安全地住在柏林的克雷恩。

1872年，李获得了他的博士学位。挪威学术界被李的学术研究所震撼，同年他们在克里斯蒂安尼亚大学为李提供了一个职位。他和他的老师西罗开始着手整理数学家阿贝尔的学术研究成果。1874年，李迎娶了安娜·伯奇，随后生育了三个孩子。

至此，李把他的学术研究集中于一个他认为已经足够成熟的题目上。在数学里，有不同种类的方程，但有两类方程是很重要的。第一类就是代数方程，这种方程已经被阿贝尔和加洛瓦有效地研究过了。另一种就是由牛顿在他关于自然定律的研究中引入的微分方程。这种方程包括从微积分中引入的概念，与直接使用物理量不同的是，这种方程描述的是物理量随时间发生的变化。更精确点说是研究量的变化率。举例来说，牛顿最重要的运动法则说，物体的加速度与所有作用在该物体上的力成比例。加速度就是速度的变化率。与直接告诉我们物体的速度是多少不同，这个法则告诉我们速度的变化率。类似地，

牛顿发明的另外一个方程是温度的变化率与物体的温度同周围温度的差成比例。

物理学中的很多重要方程都是微分方程。比如流体流动、重力作用、星球运转、热力变化、波浪运动、磁力作用、光和声音的传播等。正如牛顿最先认识到的那样，如果观察量的变化率而不是考虑量本身的话，很多自然现象都将变得相对简单并容易研究。

李给他自己提出了一个非常重要的问题。那就是是否存在一个微分方程的理论与加洛瓦的代数理论类似？是否存在某种方式来判定什么时候某个微分方程可以通过某种方法来求解？

问题的关键再次归结到了对称上。李认识到，他的几何学中的某些结果可以重新通过微分方程的术语来进行理解。给定某个微分方程的解，李可以通过某种变换（某种特殊的群）然后证明结果仍然是原方程的解。从方程的一个解可以得到多个解，所有这些解都通过群相联系。

这给很多等待我们去发现的美丽关系提供了线索。回顾一下加洛瓦对对称性的应用给代数方程带来了什么，可以设想更加重要的微分方程将会起到同样的作用！

*

加洛瓦研究的群的元素都是有穷的，也就是群内的变化数量是一个数字。举例来说，五次方程的根的所有排列的群具有 120 个元素。然而，还有一些有意义的群是无穷的，包括微分方程的对称群。

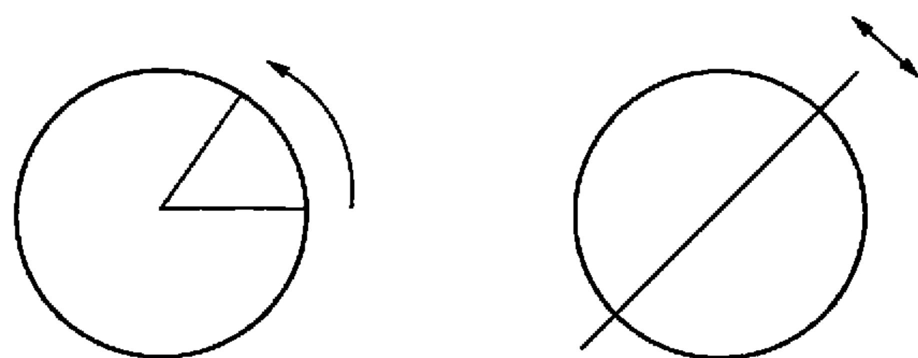
一个常见的无穷群就是圆周的对称群，这个群包含一些沿着任意角度进行旋转的变换。因为存在许多可能的角度，因此，这个圆周

的旋转群就是无穷的。这个群的表示符号是SO（2）。这里“O”代表“正交的”（orthogonal），意味着这些变化是坐标系的严格运动。而“S”意味着“特殊的”（special），旋转并不使得坐标系翻转。

圆周也具有无穷多个反射性对称的轴。如果你反射一个具有任意直径的圆，都可以得到同样的圆。加上反射性的话，我们可以得到一个更大的群，O（2）。

群SO（2）和O（2）是无穷的，但是却是无穷群中较为温和的。可以通过确定单个数字也就是一个相关的角度来决定不同的旋转。把两个旋转复合，也就是加上一个相关的角度而已。李在他的术语中把这种情况叫做“连续的”，据此可以判断SO（2）是一个连续群。因为去确定一个角度只需要一个数字就可以，SO（2）是一维的。对于O（2）也是一样的，因为我们需要做的仅仅是通过旋转来区分反射，这在几何中反应在加号或者减号上。

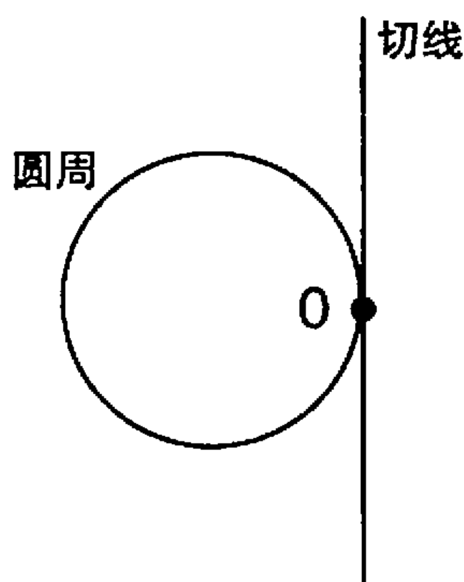
群SO(2）是李群中最简单的例子，而这个群同时具有两类结构：它是一个群，但同时也是一个簇，即一个多维空间。对SO（2）来说，簇是一个圆周，而把圆周上的两个点结合起来的群操作也就是把两个相应的角度相加。



圆周具有无穷多个旋转性对称（左），和无穷多个反射性对称（右）

李发现了李群的一个漂亮的特点：群结构可以线性化。也就是说，基本的曲线簇可以被一个平面欧几里得空间所取代。这个空间是

簇的正切空间。对 SO（2）来说，可以表现如下：



从李群到李代数：圆周的正切空间

通过这种方式进行线性化后的群结构就给了正切空间一个代数结构，而这是群结构的一个“极小的”版本，描述了变换是如何地接近恒等式的。这被叫做群的李代数。它和群具有相同的维数，但是它的几何形式却是很简单的平面形式。

可以证明，对李代数进行普通的代数操作不是乘积 AB，而是 $AB-BA$，这被称作换位（Commutator），也是李的一个深刻洞见。而对于像 SO（2）这样的群来说，$AB=BA$，换位子是 0。但是对于像 SO（3）这样的三维旋转群来说，除非 A 和 B 的旋转轴相同或者是旋转一个直角，$AB-BA$ 就不会为 0。所以群的几何在换位子的作用中才能得以显现。

20 世纪 90 年代初，“微分域”的理论被创造出来以后，李的“加洛瓦理论”是某种微分方程的梦想也变成了现实。但是事实证明，李群理论比李预期的更加重要，也应用得更加广泛。除作为一种判别微分方程是否有解的工具，李群理论和李代数已经广泛地拓展到了数学的各个领域。“李理论”已经脱离了其创造者，并且比想像的更加伟大。

事后证明，主要原因在于对称。对称已经广泛地融入数学的各个领域，并且成为数学物理学的基础。对称描述了世界潜在的规律性，而这正是推动物理学的动力。诸如旋转等连续的对称与自然、时间和物质的本质极其接近，它们意味着多重形式的守恒定律，表明一个封闭的系统既不能获得能量，也不会失去能量。这种联系是希尔伯特的学生诺特（Emmy Noether）发现的。

当然，下一步就是去尽可能地理解李群，就像加洛瓦和他的后继者从他的有穷群中推演出很多性质一样。另一位数学家加入到了这个行列中。

*

安娜很担心他的儿子。

她的医生告诉她，她的儿子威廉“非常虚弱，很笨拙”，而且“总是激动，是一个完全不切实际的书虫”。随着威廉逐渐长大，他的健康状况有所好转，但是他的书虫倾向却没有任何改变。大概在他 39 岁生日之前，他发表了一篇数学论文，被公认为是“有史以来最伟大的数学论文”。这种说法肯定有点主观，但毫无疑问的是，人人都应该去读一下威廉的论文。

威廉·卡尔·约瑟夫·基令（Wilhelm Karl Joseph Killing）是约瑟夫·基令和安娜的儿子。他有一个名叫卡尔的哥哥和一个名叫海德维格的妹妹。约瑟夫是一个法律书记员，而安娜是一个药剂师的女儿。他们在德国中部靠近东边的城市伯巴赫（Burbach）结婚，随后举家迁到了梅德巴赫（Medebach），在那里约瑟夫当上了市长，之后他又当过温特堡（Wintberg）和鲁森 (Rüthen) 的市长。

因为家庭条件比较优越，家里可以为威廉请一位家庭教师，为威廉进入体育馆做准备。体育馆在多特蒙德西边50英里左右的布里隆。上学期间，威廉非常喜欢古典学，比如拉丁文、希伯来文和古希腊文。一个名叫哈尼斯奇马切尔（Harnischmacher）的老师把威廉引向了数学。事实证明威廉非常擅长几何，并决心当一名数学家。他考入了现在明斯特的威斯特法伦·威廉大学（Westphalian Wilhelm University of Münster），当时那只是一所皇家学院。学院并不传授什么先进的数学知识，所以基令只好自学。他读了普吕克的几何学著作，并努力从中得出属于自己的新定理；他还读了高斯的《算术研究》。

两年后，他离开皇家学院去了柏林大学，那儿的数学教学质量要好得多，并且，魏尔斯特拉斯、库默尔，以及验证了能量守恒与对称的关系的数学物理学家赫尔曼·冯·亥姆霍兹（Hermann von Helmholtz），都在那儿。他以魏尔斯特拉斯的思想为基础，写了一篇几何曲面方面的博士论文，以此获得了数学和物理学的教职，另外还教授希腊语和拉丁语。

1875年，他与一个音乐教师的女儿安娜·考密尔结了婚。他们的头两个孩子都是男孩，但都夭折了；接下来的两个是女儿，分别叫玛利亚和安卡，都长大成人；后来，基令又当上了两个男孩的父亲。

1878年，他回到了自己原来的学校，但这次，他成了老师。他的工作任务很重，每周有36小时的课，但他还是挤出时间来进行数学研究——伟大人物总是这样的。他还在顶级的期刊上发表了一系列的文章。

1882年，魏尔斯特拉斯为他在布劳恩斯堡的荷西安教会学校（Lyceum Hosianum）安排了一个教授职位，他在那儿干了十年。布劳恩斯堡没有强大的数学传统，也没有可以一起切磋数学研究的同事，

但是基令好像并不需要这些东西。他在那儿作出了整个数学中最重要的发现之一，而这个发现却让他大失所望。

他的学术野心非常大：李群的完全分类。教会学校并未购进刊载有李的文章的期刊，所以基令对李的研究所知甚少。但是，1884年，他独自发现了李代数的重要性。基令了解到，每个李群都与李代数有关，而且他很快发现，李代数可能比李群更容易处理。因此，他的研究重心转到了李代数的分类上。

事实证明这个问题非常复杂——我们现在知道这个问题很可能没有一个合理的答案，以一种简明的步骤就可作出能够表示所有李代数的作图法是不可能存在的。所以基令就从一些更细小的东西着手研究，即对构成李代数的基本构成单元进行分类。这有点像想要对所有的建筑进行分类，却先去研究砖的形状和尺寸。

这种基本的构成单元就是“单李代数”。他和加洛瓦的单群理论有一个非常相似的特点，那就是不可再分为正规子群。事实上，一个单李群有一个李代数，反过来说也一样正确。令人吃惊的是，基令成功地列出了所有可能的单李代数——数学家们把这一定理叫做“分类”(classification)。

在基令看来，分类只是某种更普遍的东西的一个很局限的版本，而且，为了达到某种目的而不得不作出的种种限制性假设让他感到颓丧。他尤其厌恶缺之不可的简单化假设，这让他只能从复数转入对李代数的研究，而不是从实数开始。前者在研究中比较好用，但是与几何问题的联系却不如实数直接。由于基令个人的原因，他觉得这项研究不值得发表。

他曾试图联系李，但是未能如愿。他写信给克雷恩，克雷恩让他和李的助手弗里德里希·恩格尔建立了联系，然后就去了克里斯蒂安

尼亚大学。很快，恩格尔就和基令相处甚洽，他坚定地支持基令的研究，帮助他解决了一些技术问题，并鼓励他做进一步的研究。如果不是恩格尔，基令可能已经放弃了。

起初，基令认为自己已经知道了所有的单李群系列，而这只是两个无限李群族 so（n）和 su（n）中的一部分。SO（n）是一个包含所有 n 维空间中的旋转的正交群；而与之相似的 SU（n）则是复合 n 维空间的酉群。历史学家托马斯·霍金斯（Thomas Hawkins）曾想像“恩格尔读到基令大胆的猜想时会多么吃惊。在偏僻的东普鲁士，一所旨在培养办事员的教会学校里，有一个籍籍无名的教授，在研究李的群转换理论，还要为自己得出的理论和猜想出来的定理著书立说。”

1886 年夏天，基令在莱比锡见到了恩格尔和李，这时二人都在莱比锡工作。但是很不幸，李和基令之间存在着某种意见分歧，他始终都不太欣赏基令的研究，还总是试图贬低它的价值。

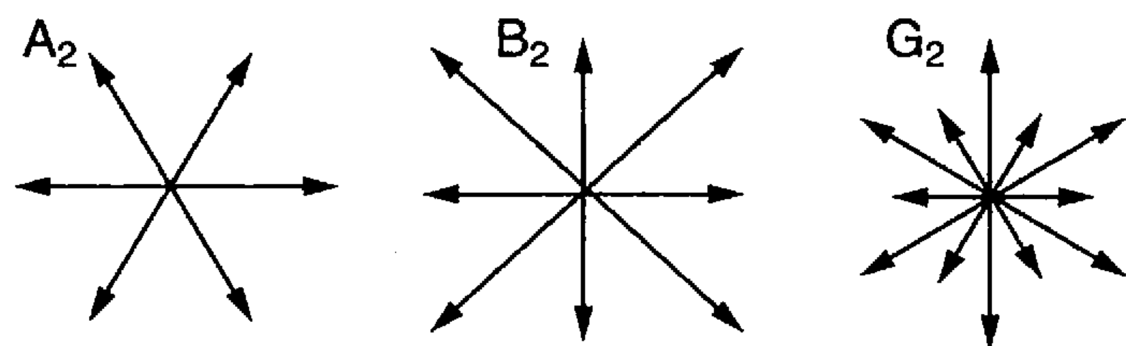
*

基令很快就发现，自己关于单李群的新颖猜想是错误的，因为他又有了新的猜想。这个新猜想对应的李群现在被称为 G_2 群，这个群有 14 维，既不像特殊的线性群，也不像正交李代数，它不属于某个族，而是个单独的例外。

如果说这很陌生，那基令在 1887 年冬天完成的最终分类就是个陌生人。在李代数原有的两个无限族的基础上，基令加上了第三个族，sp（2n），也就是我们现在所说的辛群（symplectic group）SP（2n）（今天，我们把正交群分成了两个不同的子家族，它们都适用于偶维空间和奇维空间，从而产生了四个族。这样做自有其相应的道理）。

现在，例外群 G_2 也获得了五个伙伴：56 维的有两个（其中那个较短的族渐渐不用了），78 维、133 维和 248 维的各有一个。

基令的分类是通过一个冗长的代数论证完成的，这一论证引出了几何学中一个美丽的问题。从一个假设的单李群开始，他设想出一个以多维空间中的点组成的图形，也就是我们现在所说的“根系”（root system）。对于三种单李群来说，根系存在于二维空间中。他们的根系是这样的：



二维空间中的根系

这些图形有很多的对称。事实上，这很容易让人想起万花筒中看到的图形——以一定角度组合起来的两面镜子反射出的多重影像。这些相似的图形不会重复，因为根系有自己美妙的对称群。这就是我们现在所说的外尔群（Weyl group，但这是不公平的，因为它们的发明者是基令），它们类似于我们在万花筒中看到的物体的反影。

如果能恰当地将李代数分成若干个结构类似于 su（n）的区块，那么所有关于单李代数的研究都是可以完成的，这就是基令证明的潜在框架。随后，这种分类引出了这些区块的几何问题，也就是如何运用这些美妙的对称。在对关于这些区块的几何问题进行分类整理以后，我们就可以回到我们最初想要解决的问题了——找到所有可能存在的李代数。

正像基令所说，“一个单根系对应一个单群。反过来也可以说，

一个单群是由一个单根系决定的。这样，单群的问题就解决了。因为每个 l 都有四种结构，另外再加上 $l=2$，4，6，7，8 这个例外单群。”

这里的“群”是“无限群”的缩略形式，我们把它叫做李代数，而 l 就是根系的维数。

基令提到的四种结构就是李代数中的 su（n），so（n），so（$2n+1$）和 sp（$2n$），分别对应于群族 SU（n），SO（n），SO（$2n+1$）和 Sp（$2n$），即酉群、偶维空间的正交群、奇维空间的正交群和偶维空间的辛群。辛群是不确定的位置—动量（position-momentum）的对称，位置—动量概念是汉密尔顿在自己的力学公式中引入的。辛群的维数总是偶数，因为变化的位置—动量总是成对的。基令认为，除了这四个族，还有另外六种单李代数。

他的说法大体上是正确的。1894 年，法国几何学家艾利·卡当（Élie Cartan）发现，基令那两个 56 维代数事实上是同一个代数，只不过是用了两种不同的方式去理解罢了。这也就是说，只有五种例外单李代数，它们分别对应于五个例外单李群：基令的老朋友 G_2，另外四个分别是 F_4，E_4，E_7 和 E_8。

这是一个极其奇怪的答案。无限族已经相当足够了，它们都与众多不同维度的几何形态相联。但是这五个例外李群看起来和几何学毫无关系，而其维度也很古怪。为什么 14 维、56 维、78 维、133 维和 248 维是特殊的呢？这几个数字有什么不同寻常之处呢？

这个有点像想要列出所有砖的形状时，找到了类似如下的答案：

- 大小为 1，2，3，4……的长方形的砖。
- 大小为 1，2，3，4……的正方形的砖。
- 大小为 1，2，3，4……的薄板砖。

- 大小为 1，2，3，4……的棱锥形砖。

答案也可能很简洁，就像下面这一串：

- 大小为 14 的四面体。
- 大小为 52 的八面体。
- 大小为 78 的十二面体。
- 大小为 133 的十二面体。
- 大小为 248 的十二面体。

就这么些，再没有其他的了。

为什么会有形状和大小都这么奇怪的砖呢？这些砖有什么用呢？

这看起来完全是疯了。

看起来是疯了，事实上，基令对这些例外李群的存在感到非常沮丧，曾一度希望这只是一个有朝一日可以连根拔除的错误。基令认为它们破坏了自己分类的美感。但是它们就这么存在着，而最后我们还是理解了它们存在的原因。从很多角度看，现在这五个例外李群都比那四个无限群有趣。它们对粒子物理学将会很重要，也许将来我们就会发现；而他们对数学的重要性则非常明显。它们还有一个未被充分揭示的神秘统一性，这种统一性把它们与汉密尔顿的四元数联系起来，甚至从更广泛的意义上说，它们还与八元数联系了起来，这实在让人吃惊。

这一系列的思想都非常美妙，而且都是基令一个人的发现。当然，他的研究中出现了一些错误，有些证明不太说得通。但是这些错误很久以前就已经被纠正过来了。

*

最伟大的数学文献的命运总是如此。基令那个时代的人是怎么看待他的著作的呢？

对此我们知道的并不多。李的嘲弄对基令的伟大著作肯定没有好处。不知道出于什么原因，两个人闹翻了，在李看来，基令没有取得任何重要的成果。更糟糕的是，这也许就是李最想自己去证明的定理。受到了这样的打击，他就只能吃不到葡萄而说葡萄酸了。李说，在这一领域，只要是自己没有做的，都是垃圾。不过，他没有说得这么露骨。

这让基令难以对自己的定理的价值作出正确的评价。对基令来说，他所取得的成就只是某个重要问题的不起眼的小配角，而对所有的李群进行分类才是重要的，可自己却没有做到。基令是一个低调的人，而李却尽其所能地让基令的姿态变得更低。

从任何方面讲，基令都是超前的。很少有数学家看到基令的理论将会变得多么重要。他们顶多将其看成与微分方程有关的几何学的一个技术性分支。

最后，基令变成了一个极具责任感和人道主义精神的虔诚的天主教徒。他以阿西西的圣方济各为榜样，并在39岁时和妻子一起加入了方济各会的第三会（Third Order）。他仿佛已经变成一个不知疲倦地为自己的学生工作的体面男人了。他是一个保守的爱国主义者，第一次世界大战后四分五裂的德国社会让他悲痛万分；而两个分别死于1910年和1918年的儿子让基令的情绪越来越糟糕。

基令著作的真正价值是在1894年被人们发现的。当时，卡当在自己的博士论文中重新得出了完整的理论，不仅在单李群的分类问题

上前进了一大步，还用矩阵语言将其表述了出来。卡当直言不讳地将自己差不多全部的思想都归功于基令。他对基令的所有成果都进行了整理，补充了一些漏洞（有些是很严重的漏洞），还更新了一些术语。但是有一种传闻迅速被夸大，说基令的理论充满漏洞，功劳应该属于卡当。很少有数学家是合格的历史学家，他们总是褒扬自己知道的著作，而不去赞扬促成这些著作问世的更早的作品。所以基令的很多思想都被归在了卡当的名下。

只要是读过基令著作的人就很快会发现传闻是错误的。这些著作中的思想条理清晰，其中的证明有些过时，但几乎完全正确。更重要的是，书中所用的思想清一色是为了一个预期的结果而巧妙地选择出来的。这是最高级别的数学，没有人能做到如此尽善尽美。

但是很不幸，几乎没有人读过基令的著作。他们只读过卡当，所以基令的功劳被他们忽略了。但是最终，基令的著作得到了应有的认可。1900 年，他获得了卡赞物理数学协会（Kazan Physico-Mathematical Society）颁授的罗巴切夫斯基奖（Lobachevsky Prize）。这是这个奖项的第二次颁授，而第一次获奖的就是李。

基令死于 1923 年。直到今天，真正知道这个名字的人仍没有应该知道这个名字的人多。他是有史以来最伟大的数学家之一，无论如何，他留给我们的遗产都是不朽的。

专利局职员

20 世纪初，群论开始进入基础物理学，它在这一领域引起的变革和在数学中引起的一样重大。

在 1905 年这个黄金年份，一个将要成为那个时代最具标志性的科学家的人发表了三篇论文，每一篇都革命性地改变了一个独立的物理学分支。他那时既不是教授也不是科学家。他上过大学，但是却没能获得教职，而是瑞士伯尔尼一个专利局的一个办事员。没错，他的名字叫阿尔伯特·爱因斯坦。

如果说有人能够象征 20 世纪的物理学，那这个人就是爱因斯坦。另外，他还是数学天才的象征，事实上，他仅仅是一个数学能力出众的数学家而已，而没有加洛瓦和基令那样的创造力。爱因斯坦的创造力没有体现在创造新型的数学上，而是体现在精准、直观地把握物理世界的超人能力上。在物理学领域，他出色地运用了既有的数学成果。爱因斯坦还有着出众的哲学天赋。他从最简单的原理中抽出了精妙的理论，与仰仗通过实验获取广博的知识相比，他更信奉美感。他认为，

总能从重要的观察中总结出关键性的原理。美就是通往真的大门。

关于爱因斯坦的生活与工作的著述和学术研究已经浩如烟海，仅仅一章不可能全面、深入地阐释爱因斯坦。但他是对称理论发展史上的关键人物：不是别人，正是爱因斯坦，发起了将对称数学引入基础物理学的一连串事件。但我认为爱因斯坦并不这么看，对他来说，数学只是物理学的一个侍从，而且常常是不太顺从的一个。只是到后来，沿着爱因斯坦留下的记号，及其前期开拓道路时留下的杂乱踪迹，后人才得以揭示作为其思想基础的漂亮的数学观念。

因此，我们必须重新看待这个不起眼的专利局职员陡增的声望——一个技术层面上的三流专家，准确地说，根本就是不入流。因为他只是整个故事的一部分，所以我只想讲一下与之相关的事件。如果你想更深入、更全面地了解爱因斯坦的生平，可以读一下亚伯拉罕·派斯（Abraham Pais）的《上帝是不可捉摸的》(*Subtle Is the Lord*)。

是的，是不可捉摸的，而不是像爱因斯坦曾经认为的：恶意的。

爱因斯坦几乎对宗教没有任何兴趣，一生都致力于那些解释宇宙的原理，而这些原理都是沿着数学的轨道发展的。他最著名的那些言论都援引了神性，但只是把它当做一种宇宙秩序的象征罢了，而不是一种关切人类事物的超自然力量。他信仰的是一种无神的、非宗教性的仪式。

*

爱因斯坦通常被视为牛顿的天然继承人。之前的数学家已经对牛顿的“世界体系”(System of the world，他的《自然哲学的数学原

理》的副标题）进行了一些补充，但是爱因斯坦是第一个让这一体系发生实质性变革的人。最重要的早期理论家是麦克斯韦（James Clark Maxwell），他的电磁学方程解释了电磁现象，尤其是光的现象，还没有超出牛顿理论体系的范围。爱因斯坦走得更远，他作出了重大变革。具有讽刺意味的是，这种修正了万有引力理论的变革是从麦克斯韦的电磁波理论（光及其相对性）发展出来的。更加讽刺的是，这一理论的基本特征是，光波发挥着关键作用，但是牛顿认为光不可能是波；最重要的是，第一个证明光是波的实验其实是由牛顿完成的。

科学家对光的兴趣至少可以追溯到亚里士多德，尽管亚里士多德事实上是一个哲学家，总是问一些科学家们才会问的问题。我们是怎么看到东西的呢？亚里士多德认为，当我们注视一个物体，物体就对自身和人眼之间的介质（我们现在把这种介质叫做“空气”）发生作用。然后，眼睛就觉察到了介质的这种变化，然后就有了视觉。

这种解释在中古时期被推翻了。当时的人们认为，我们的眼睛放射出一种射线，照亮了我们注视的物体。不是物体把信号传递给眼睛，而是眼睛的视线覆盖整个物体。

事实上，人们已经了解到，我们是通过反射光看见物体的，在日常生活中，太阳光是主要的光源。实验显示，光是以“射线”的形式直线传播的。当光线遇到合适的表面被弹出时就会发生反射。太阳光的射线投射在所有未被遮挡的物体上，这些光被弹向各个地方，有一些光线进入了人眼，人眼就从这个方向上接收到信号，大脑对这些人眼接收到的信号进行加工，于是我们就看到了反射光线的物体。

主要的问题是，光是什么？围绕着光有很多的谜。它不只会反射，还会折射——在通过两种不同的介质（比如空气和水）时，它的传播方向会发生改变。这就是为什么插入水池的木棒看起来是弯的，

这也是眼镜的工作原理。

更令人费解的是光的衍射现象。1664年，科学家兼博物学家罗伯特·胡克（Robert Hooke，他的研究总是与牛顿撞车）发现，如果把一个透镜放在一个平镜上，然后透过透镜进行观察，就会看到若干同心的彩色环。这些环现在被称为“牛顿环”，因为牛顿是第一个分析其构成的人。现在，我们认为这个实验证明了光就是一种波：这些环就是干涉条纹，在这些条纹上，波与波有时会相互抵消。但是牛顿不认为光就是波。由于光是沿直线传播的，所以他认为光一定是某种粒子流。他在完成于1705年的《光学》中说：“光是由发光体发射出的粒子或光子组成的。”粒子论可以很容易地解释反射现象：粒子撞上某种表面时弹起。但是这种理论却很难解释折射现象，而面对衍射现象时更是土崩瓦解。

在思考光的弯曲现象时，牛顿断定介质才是问题的根源，而不是光本身。这使他认为存在一种能比光更快地传播振动的“以太”(aethereal medium)。他相信辐射热就是这种振动的很好的例证，因为他已经坚信热辐射可以穿过真空。所有真空中的东西都一定带有热，并形成折射和衍射。用牛顿的话说就是：

> 发热体发出的热难道不是通过空气被抽走后留下来的、比空气更敏锐的介质，以振动的形式在真空中传播的吗？它和造成光的折射与衍射的介质难道不是同一种介质吗？难道不也是通过它的振动，光才得以把热传给物体，并使其易于反射和传播吗？

读到这段文字，我禁不住想起我的朋友特里·普莱切特（Terry Pratchett），他的系列小说以“荒诞世界”为背景讽刺了我们的现实

世界。他书中的术士、女巫、浪游者以及小矮人都爱拿人类的缺点开涮。在荒诞世界中，光的传播速度和声音差不多，这样就可以看清黎明一步步穿越大地。黑暗是黎明必需的对立面，而黑暗却比光走得要快，因为它要给光让道。即使在我们的世界中，这也都是说得通的，但是事实令人失望，这些东西全都是不正确的。

牛顿关于光的理论也有同样的缺点。牛顿并不傻：他的理论看起来回答了很多重要问题。但是可惜，这些答案都建立在错误的基础上：他认为辐射的热和光是两种不同的东西。他认为光遇到一个表面时，就会引起热的振动，而那些导致了光的衍射现象的振动也是这种振动的变种。

因此，通过非常牵强的证明，“光以太”（luminiferous ether）的概念诞生了。其实在人们认识到光就是波以后，为波提供介质的正是以太（现在，我们认为光既不是粒子也不是波，而是有点同时像这二者的东西——波粒子。但现在讲这些还太早）。

但是什么是以太呢？牛顿的说法很简单：“我不知道什么是以太。”他认为，如果以太也是粒子构成的，那这些粒子一定比空气的粒子更轻、更小，甚至比光的粒子更轻、更小——就像在荒诞世界中一样，这是为了给光让道。牛顿说，“粒子的极小，”“使其互相避让的能力变得极大，这就使介质变得比空气稀薄得多，也有弹性得多。这样，它们就几乎一点也阻碍不了抛射体，也更能通过膨胀自己来挤压粗糙的物体。”

在之前的1678年，荷兰物理学家克里斯蒂安·惠更斯提出了不同的理论：光是一种波。这一理论清楚地解释了反射、折射和衍射——我们可以看到与之类似的效果，比如水波。以太之于光正如水之于海浪，当波经过时，他们就会产生运动。但是牛顿不同意这种看法。这

场论战非常混乱，因为两位科学家都对波的性质提出了错误的假设。

麦克斯韦介入论战后，一切都随之改变了，但是，能取得这样的成果，是因为他站在了两位巨人的肩膀上。

*

电热、电灯、收音机、电视机、食物加工机、微波炉、电冰箱、吸尘器，以及数不尽的工业产品都源自一个人的洞见，这个人就是米歇尔·法拉第。1791 年，法拉第出生于伦敦的纽英顿·波特（Newington Butts，也就是现在的大象堡）。他是一个铁匠的儿子，在维多利亚时代成了著名的科学家。他的父亲是桑地马尼安教派（Sandemanian）的成员，这是基督教中的少数派。

1805 年，法拉第成了一个图书装订学徒，并开始进行科学实验，特别是化学方面的实验。1810 年，他成了“市哲学协会”（一个年轻人一起讨论科学问题的组织）的成员，大大拓展了自己对科学的兴趣。1812 年，他获得了英国首席化学家汉弗里·戴维（Humphry Davy）在皇家学会举行的最后一次讲座的赠票。此后不久，他恳求戴维给他一份工作，他参加了面试，但并没有合适的工作可干。在戴维的化学助手因挑起斗殴而被开除后，法拉第接替了他的工作。

从 1813 年到 1815 年，法拉第跟着戴维和戴维的妻子在欧洲各地游历。拿破仑给了戴维一张通行证，并允许他带一名侍从，所以法拉第扮演了这一角色。令法拉第气愤的是，戴维的妻子珍妮望文生义，真拿他当自己的仆人使唤。1821 年，法拉第的遭际更趋好转：他得到了提升，还和一个著名的桑地马尼安会士的女儿、莎拉·巴纳德（Sarah Barnard）结了婚。但是，他的电磁学研究才刚刚起步。追随着

丹麦科学家汉斯·奥斯特（Hans Ørsted）的脚步，法拉第发现，当电通过磁铁附近的线圈时会产生一种力。这就是电机的原理基础。

然后，他的研究兴趣被繁重的管理和教学事务淹没了，但这也给他带来了非常有利的影响。1826年，他开始举行关于科学的夜间讨论，并为年轻人开办了圣诞讲座，现在这两项活动还在开展。现在，圣诞讲座还被搬上了电视，而电视也正是法拉第的发现的产物。1831年，法拉第终于回到了他的实验上，他发现了电磁感应。通过实验，他发现电产生的力肯定是某种粒子，而不是通常被认为的流。

辉煌的科学成果往往会把科学家推上行政职位，也会进而扼杀这门科学的活力。法拉第奉命为负责海上航路安全的英国港务局提供科学指导。他发明了一种性能更好的油灯，使它发出的光更亮。1840年，他当上了桑地马尼安教派的长老，身体状况却开始恶化。1858年，他获得了汉普顿宫一间“公馆式”住宅的免费居住权，这幢建筑原来是亨利八世的宫殿。他死于1867年，葬在海格特公墓（Highgate Cemetry）。

*

法拉第的发明革新了整个维多利亚世界，但是也许由于早期受教育不足，他并不长于理论，而他解释自己的发明的工作原理时，总要借助古怪而呆板的类比。就在1831年法拉第发现电磁转换原理的时候，一个苏格兰律师带着自己的儿子（这也是他惟一的孩子）来了。这位律师更关心自己土地的所有权，也很想教育好小“贾麦西”(Jamesie)，孩子正式的名字是詹姆斯·科莱克·麦克斯韦。

天资聪颖的贾麦西对机器很着迷。“它是怎么弄的啊？”这是他

的典型问题：它是如何运转的？另一个经典问题是，“是什么让它这样运转的啊？”他父亲对这些东西也很好奇，就尽其所能地向他解释。如果父亲解释得不够深入，贾麦西就会紧跟着问一句：“到底是什么让它这样运转的啊？”

詹姆斯的母亲在他 9 岁时死于癌症，这更拉近了父子俩的关系。孩子被送进了爱丁堡学院，这所学校分科很细，要求学生整洁得体，在规定学科上很有实力，但循规蹈矩的教学方法导致其完全缺乏创造性思想。贾麦西并不是老师们喜欢的那种学生，而他父亲为他设计的奇装异服也违背了学校对整洁的要求。于是别的孩子就为詹姆斯起了个“呆瓜”（Dafty）的绰号。尽管孩子们理解不了詹姆斯，但是他的顽强还是赢得了他们的尊重。

学校给詹姆斯带来了一样好东西，那就是对数学的兴趣。他在一封给父亲的信中还在谈论如何做“一个四面体、一个十二面体，还有两个我还不知道名字的物体（他说的可能就是八面体和二十面体）”。14 岁时，他因独立发明了一个现在被称为笛卡尔卵形曲线的数学曲线而获奖，因为其最初发明者是笛卡尔。他的文章还被递交给了爱丁堡的皇家协会。

詹姆斯也写诗，但他的数学天赋更高。16 岁时，他进入了爱丁堡大学，后又进入英国最好的数学研究机构、剑桥大学继续深造。指导他考试的威廉·霍普金斯说：“他是我见过的最卓尔不群的人。”

取得学位后，他继续留在剑桥大学攻读研究生课程，做关于光学的实验。后来，他读了法拉第的《实验研究》(*Experimental Researches*)，并开始研究电学。长话短说，他把法拉第的电磁现象做成了模型，并在 1864 年将其纳入到一个由四个数学法则（用当时的标记方法，这些法则不止四种，现在我们可以用向量标记法将其分为

四种，有些形式主义者甚至将它们归结为一种）构成的体系当中。这种法则把电和磁描述成两个场，一个属于电，一个属于磁，这样就涵盖了整个空间。这两个场不仅描述了各个方位上的电和磁的强度，还描述了二者的方向。

这四个方程的物理学意义很简单。其中两个告诉我们电和磁既不能被创造，也不会被毁坏；第三个方程讲的是不断变化着的磁场对周围电场的影响，并赋予了法拉第的发现以数学的形式。第四个方程讲的则是不断变化的电场对周围磁场的影响。就算是用文字表述，这些方程也很有美感。

通过这四个方程进行简单的推导之后，麦克斯韦证实了自己的猜测：光是一种电磁波，其传播会对电场和磁场产生干扰。

根据数学的逻辑，所有的数学家都可能会发现，从麦克斯韦的方程很容易得出某种东西——“波方程”。看其名称就可以知道，它是描述波的传播的。麦克斯韦的方程也对波的传播速度作出了推测：它一定是以光速传播的。

只有一种东西可以以光速传播。

当时的人们猜想，波只可能是存在于某种东西之内的波。它必定有其赖以传播的介质，而波就是这些介质的振动。最明显的例子是，光波就是以太。数学告诉我们，波的振动必定与传播方向构成一定的角度。这也就揭示了牛顿和惠更斯的困惑：他们认为光一定是沿着传播的方向振动的。

麦克斯韦的这一理论还作出了另一个推测：电磁辐射的波长（一波与临波的距离）可能是多种多样的。光的波长很短，但可能还有波长比它长得多的波存在。这一理论还启发了海因里希·赫兹（Heinrich Hertz），使他发明了我们现在所说的无线电波。紧接着，古里尔莫·马

可尼（Gulielmo Marconi）就发明了送话器和听筒，很快，我们就几乎可以在整个地球的范围内进行即时通话了。现在，我们还通过这种方式传送图片，通过雷达进行空中监测，通过全球定位系统进行导航。

但是，这就使以太的概念成了一个难题。如果以太存在，那么围绕太阳旋转的地球就一定会受制于它。如果它存在，我们应该可以探测到这种运动；要么，我们就必须通过实验证实这种观念的错误。

对这一谜题的回答将会彻底改变物理学的面貌。

*

1876年夏天，位于符腾堡公国的乌尔姆市，由两个犹太商人创办的伊兹雷尔和列维（Israel and Levi）公司，加入了一个新的合伙人，赫尔曼·爱因斯坦（Hermann Einstein）。赫尔曼在年轻时就已经显示出了相当的数学天赋，可是他的父母却供不起他上大学。现在他变成了一个出售羽绒床垫的公司的合伙人。

八月，赫尔曼与保琳·科克（Pauline Koch）在坎施塔特（Cannstadt）的犹太教堂里结了婚，夫妇俩在班霍夫街（Bahnhofstrasse）——也就是该市的车站路——安了家。过了不到八个月，他们的第一个孩子就出世了。出生证明上写道："一个男孩，取名阿尔伯特，生于乌尔姆，住在（赫尔曼）家，其妻保琳·爱因斯坦·科克，犹太人。"五年后，阿尔伯特有了一个妹妹，名叫玛利亚，两人的关系非常亲密。

阿尔伯特的父母都对宗教抱持着宽松的态度，努力融入当地文化。当时，很多的德国犹太人都是"同化主义者"（assimilationist），他们淡化自己的文化传统，以期更好地适应周围居民的信仰。赫尔曼和保琳给孩子取的名字并不是传统的犹太人的名字，但是他们说这是

随了孩子的祖父“亚伯拉罕”（Abraham）的名字。赫尔曼家很少谈论信仰问题，阿尔伯特也不遵循犹太人的礼数。

玛利亚发表于1924年的童年回忆，是我们了解阿尔伯特的早期经历和性格的主要资料来源。很明显，他出生时吓了母亲一跳，他的后脑勺异常尖削，而且出奇地大。“太重了，太重了！”她第一次看到孩子的时候，叫了起来。这孩子学说话费了很长时间，以至于大家都担心他将来会是个智障。但这只是因为阿尔伯特只有在自己完全想明白之后才会去做一件事。他后来说，只有当自己能够组织起完整的句子时才愿意开口说话。他会在头脑里一遍一遍地尝试，在最终确定自己的语言准确无误的时候，才会说出来。

阿尔伯特的母亲是个颇有建树的钢琴家。6岁到13岁之间，阿尔伯特跟着一个叫施密德（Schmied）的人学拉小提琴。他后来非常热爱小提琴，但是童年时却很讨厌上提琴课。

羽绒床垫的生意失败后，赫尔曼和他的兄弟雅各布合伙，做起了供应煤气和水的生意。雅各布是一个工程师兼实业家，兄弟两人在这项冒险中投入了巨额资金。后来，雅各布决定兼顾电的生意，他们经营发电设备，而不是电器产品。在保琳的父亲和其他家庭成员的资金支持下，两兄弟的公司在1855年步入正轨，而且两人还搬入了位于慕尼黑的同一所房子。起初，公司运营良好，爱因斯坦电力公司（Elektronische Fabrik J. Einstein und Co.）在慕尼黑地区出售发电设备，甚至还远销意大利。

爱因斯坦说他对物理学的兴趣是在父亲给他看了一个指南针之后被激发起来的。四五岁的时候，阿尔伯特对指南针无论怎么转动都指向同一个方向的能力特别着迷，这使他第一次窥见了物理世界之神奇。

在学校里，阿尔伯特成绩不错，但没有显示出什么特别的天赋。他慢条斯理，成绩优秀，但是不善交际。他更喜欢自娱自乐，尤其喜欢用卡片搭建房子。他不喜欢体育运动。1888 年去了体育馆后，他学好了拉丁语，直到在 15 岁离开那里的时候，他的拉丁语和数学在班里都是拔尖的。他数学方面的才能是被他的叔叔雅各布激发出来的，雅各布是个工程师，学过一点高等数学。阿尔伯特家有一个叫马克思·塔尔穆德（Max Talmud）的朋友，也对他的教育产生了重要影响。塔尔穆德是一个学医的贫困生，赫尔曼和保琳每个星期四晚上都会请他到家吃晚饭。他给了阿尔伯特几本科普书，后来他又启迪这孩子去读康德的哲学著作。他们会一连几个小时地谈论哲学和数学。塔尔穆德曾写道，自己从未见过爱因斯坦和别的孩子一起玩耍，而且他读的都是些很严肃的书，没有轻松的读物。他惟一的放松方式就是演奏音乐，其中包括贝多芬和莫扎特的奏鸣曲，保琳会为他伴奏。

1891 年，在得到一本他后来称为“几何圣经”的欧几里得几何学著作后，阿尔伯特对数学的热情爆发了。书中吸引他的主要是清晰的逻辑和欧几里得发现的思维模式。爱因斯坦一度由衷地感激学校的强制训诫（在旧教中，这是必须无条件服从的）和家庭中的犹太信仰教育。但是一旦他找到了科学，这些东西也就被弃之不顾了。自此，他希伯来文的学习和为成人礼所作的努力全都骤然中断。阿尔伯特听到了新的召唤。

1890 年代早期，爱因斯坦电力公司事事不顺，产品在德国的销售越来越难，而意大利代理人洛伦佐·加罗尼（Lirenzo Garrone）建议把公司搬到意大利。1894 年 6 月，德国的公司关张了，家里的房子也待价而沽，爱因斯坦一家也搬到了米兰，只有阿尔伯特没有去，因为他要留下来完成学业。“爱因斯坦与加罗尼”在帕维亚开设了门店，一

家人陆续地搬了过去，只有阿尔伯特留在了慕尼黑。

阿尔伯特讨厌这段令人沮丧的经历。糟糕的还不止于此：他还面临着去服兵役的危险。在未通知父母的情况下，他决定去意大利。他劝服了家庭医生，给他开了一个证明，说他有神经紊乱症状，而且很可能真有这种症状。这样，他得以提前离开学校，于1895年春天悄悄地来到了帕维亚。这吓坏了他的父母，因此他承诺继续自己的学业，以便通过苏黎世ETH（瑞士联邦工学校，现在瑞士首屈一指的高等教育机构）的入学考试。

阿尔伯特在意大利灿烂的阳光下茁壮成长。10月，他参加了ETH的入学考试，但是没有通过。他轻松地通过了数学和科学的考试，但是却栽在了人文学科上。他的论文写作好得出奇，而这也变成了进入EHT另一门径，条件是要获得高中毕业证，然后就可以自动升入EHT。于是他以文特勒（Winteler）家的寄宿客的身份进入了一所高中。文特勒家有七个孩子，阿尔伯特就加入到他们中间，和他的代父母产生了一段持久的感情。他盛赞这所学校的“自由精神”和出众的教师——直言不讳地说这些教师从不屈从外界的权威。

他有生以来第一次感受到了上学的快乐。他变得自信起来，而他的心迹也开始被别人所了解。他有一篇用法语写成的学校文章，道出了自己的未来计划，那就是研究数学和物理。

1896年，他进入了EHT，恢复了符腾堡居民的身份，但是没有国籍。他在每月的零用钱中省下五分之一，作为将来加入瑞士国籍的费用。但是他父亲和叔叔合办的电力工厂倒闭了，很大一部分家产都赔了进去。雅各布在一家大公司找了个正式的工作，但是赫尔曼打算重整旗鼓，他准备再开办一项事业。他不顾阿尔伯特的反对，又在米兰干了起来，但在苦苦维持了两年之后，还是失败了。阿尔伯特一度

因为家庭的时运不济变得更加沮丧，直到他的父亲和雅各布一样找了个安装发电站的工作才有所好转。

在EHT，阿尔伯特把很多时间都花在了到实验室做实验上。他的教授海因里希·弗里德里希·韦伯（Heinrich Friedrich Weber）对阿尔伯特的印象并不太好。“你是个聪明的孩子，爱因斯坦，你很聪明，”他对这个年轻人说，“但是你有一个很大的缺点：别人说的话你一点也听不进去。”他中止了阿尔伯特关于地球的运动是否和以太有关的实验，这个实验的目的是验证所有的流体都会传播电磁波的假设。

而爱因斯坦对韦伯也印象不佳，他觉得韦伯的课已经过时了。尤其令他失望的是，韦伯没有多给他讲一些麦克斯韦的电磁理论，所以只好自学，他读的是1894年的德语版。他还去了听著名数学家胡尔维兹和赫尔曼·闵可夫斯基（Hermann Minkowski）的数学讲座。闵可夫斯基是个很有原创性的思想家，他为数论引入了一套根本性的新方法，这些方法后来为相对论作出了巨大贡献。阿尔伯特也读了一些关于达尔文进化论的著作。

为了在EHT继续待下去，他必须获得一个助手的职位——也就是我们现在所说的助教——这样他就有钱在EHT继续深造了。韦伯曾暗示自己会给他一个这样的职位，但后来却没有了下文，因而阿尔伯特从没有彻底原谅过他。他写信问胡尔维兹能不能给他一个这样的职位，他收到的回答无疑是肯定的，但却没有最终实现。1900年底，他还没有找到工作。但是，他发表了自己的第一篇研究文章，文章讲的是分子之间的力。此后不久，他获得了瑞士国籍，他一直保留着这一国籍，直到他移民去了美国后还保留着。

整个1901年，阿尔伯特一直试图在大学里谋职，只要看到哪里有空职，就给人写信、寄送自己的文章，但是无一如愿。极度绝望的

他只得找了个高中临时教师的工作。让他吃惊的是，他发现自己竟然很喜欢教书；另外，这个工作留给了他大量的时间去继续自己的物理学研究。他写信给自己的朋友马塞尔·格罗斯曼，说自己在研究气体理论，研究物体在以太中的运动。后来，他又去了另一所学校，还是当临时教师。

格罗斯曼向阿尔伯特伸出了援手：马塞尔的父亲答应把阿尔伯特推荐给伯恩的联邦专利局的负责人。职位正式发布的时候，阿尔伯特就去应聘了。1902 年初，尽管自己还没有收到被录用的正式通知，他还是辞去了学校的教职去了伯恩，可能是他已经得到了相关消息，也可能是因为他对此很自信。1902 年 6 月，阿尔伯特正式入职了。虽然这不是他一直觊觎的学院职位，但却收入不菲，高达一年 3500 瑞士法郎，这足够他的衣食和住房之用，而且他还有充足的时间进行物理学研究。

在 EHT 时，他结识了一个名叫米列娃·马利克（Mileva Maric）的女生，她对科学有着浓厚的兴趣，而且对阿尔伯特有意。他们恋爱了。可惜保琳·爱因斯坦不喜欢她这个未来的儿媳妇，两人的感情闹得很僵。后来，赫尔曼罹患了恶性心脏病，他在弥留之际同意了阿尔伯特和米列娃的婚事，然后要求所有的家庭成员离开，以求静静地死去。阿尔伯特一生对此都有负罪感。1903 年 1 月，他和米列娃成婚，他们惟一的儿子汉斯·阿尔伯特出生于 1904 年 5 月。

专利局的工作很适合爱因斯坦，由于他出色地履行了自己的职责，1904 年底，这份工作成了他的永久性工作。但是专利局的头头警告说，以后想要晋升的话，爱因斯坦就必须去处理一些机器技术问题。借助于统计力学，他的物理学研究也获得了巨大进展。

关于让这位专利局职员获得诺贝尔奖的文章的一切，都离不开 1905 年这个“黄金之年”。在这一年，他还获得了苏黎世大学的博士

学位。此外，他还晋升为二级技术专家，获得了每年 1000 瑞士法郎的奖金——看样子那时他已经掌握了机器技术。

即使在成名之后，爱因斯坦仍然很感激格罗斯曼为他在专利局铺平的道路。爱因斯坦说，且不论其他方面，重要的是这份工作使他的物理学研究变得可能。这份完美的工作真是如有神助，爱因斯坦从未忘记。

*

在物理学历史上最重要的一年，爱因斯坦发表了三篇重要的研究文章。

有一篇是关于布朗运动的，研究的是悬浮在流体中的微粒的不规则运动。这一现象是以其发现者、植物学家罗伯特·布朗的名字命名的。1827 年，他透过显微镜观察花粉悬浊液时发现了这一现象。他注意到，在花粉的孔穴中，有一些更小的颗粒在不停地做着不规则运动。1880 年，托瓦尔德·蒂厄勒（Thorvald Thiele）作出了这种运动的数学模型，1900 年，路易·巴契里耶（Louis Bachelier）也独力发现了这一数学模型。但是巴契里耶的灵感并不是来自于布朗运动，而是股票市场的随机波动，这两种数学模型却是一致的。

而布朗运动的物理学解释仍然有待发现。爱因斯坦和波兰科学家斯莫鲁霍夫斯基（Marian Smoluchowski）都发现，布朗运动很可能就是一个当时未被证明的理论的证据，这个理论就是，物体是由原子组成的，而原子又组成了分子。根据所谓的“分子运动理论”，气体和液体中的分子不停地相互碰撞，这样就导致了不规则运动。爱因斯坦用充足的数学的推算证明了这一过程和人们观察到的布朗运动的一致性。

第二篇论文是关于光电效应的。亚历山大·贝克勒尔（Alexander Becquerel）、威洛比·史密斯（Willoughby Smith）、海因里希·赫兹及另外一些人都发现，有几种类型的金属在受到光照时会产生电流。爱因斯坦根据量子力学推测，光是由极小的粒子构成的。他的推论显示，这种假设和实验数据吻合得很好。这是量子理论第一个强有力的证据。

这两篇文章都实现了物理学中的重大突破。但是第三篇比这两篇还要厉害得多。这一篇是关于狭义相对论的，这一理论超越了牛顿的理论，大大改变了我们对于空间、时间和事物的看法。

*

我们平时都是以欧几里得和牛顿的方式观看空间的。空间有三个维度，就像一座建筑物的直角的三个方向——东、北、上。空间结构无论占据着什么样的空间，其构成形式都是一样的。物体在空间中的移动形式有如下几种：转动、被镜子反射，或者“平移”——在不转动的情况下向一侧移动。抽象一点说，我们可以认为，空间自身也可以进行同样的变换（参照系的改变）。空间结构，以及表述空间结构并作用于它的物理定律，就是这些变换中蕴含的对称。对称是在每时每处都一样发生作用的物理定律。

从牛顿物理学的观点看，时间形成了独立于空间之外的另一维度。时间只有一维，其对称变换也更简单。它可以平移（将一段固定的时间置入每一观察资料），也可以反射（时间反演——只能在思想实验中才能实现）。物理定律与你的测量数据是无关的，因此它具有时间平移对称性。大多数基本的物理定律在时间反演的情况下也是对

称的，这是一个很神秘、但又不是完全神秘的现象。

但是当数学家和物理学家们开始思考新近发现的电磁学定律的时候，牛顿物理学就失效了。使物理定律保持不变的时间和空间的变换不再是平移、旋转、反射之类简单的“运动”了；另外，这些变换也不能单独作用于时间或空间。如果你只使空间发生了改变，那么方程式就会乱成一团，必须辅以时间的改变才行。

在一定范围内，只要你所研究的对象系统是不变的，这些问题就可以忽略。但是，在关于电子移动的数学研究中，这个问题就无法回避了，同时，这也是19世纪末物理学中的问题。人们关于对称的疑虑渐渐变得严重起来。

1905年之前的几年中，不少物理学家和数学家都对麦克斯韦方程的古怪特性感到头疼。在实验室中做电磁实验和在一列行驶着的火车上做电磁实验，各自会得到什么样的结果呢？

当然，几乎没有人会在行驶的火车上做实验，但他们肯定都是在运动着的地球上工作的。出于很多原因，我们可以认为地球是静止的，因为实验设备是随着它的运动而运动的，所以它的运动并不会对实验造成影响。牛顿的运动定律，都局限于惯性参考系中，在此参考系中，运动都是匀速直线运动。地球的运动速度几乎是恒定的，但是它沿自己的轴自转，同时还绕着太阳公转，所以，相对于太阳，地球的运动是非直线运动。但是实验设备却几乎是沿着直线运动的，即使根据实验的要求调整设备的运动曲线，往往也是于事无补的。

如果麦克斯韦的方程以旋转参照系为标准，那么人们也就不会为此而感到困惑了。但他们发现了一个更头疼的问题：麦克斯韦的方程运用了另一种惯性参照系。一辆行驶着的火车上的电磁和一个固定的实验室里的电磁是不同的，即使火车匀速直线行驶，也是如此。

还有另一个麻烦：我们是可以说火车或地球是运动的，但是运动的观念是相对的。比如，我们通常注意不到地球的运动。地球的自转解释了太阳在早上升起，在晚上落下的现象，但我们并不能感觉到地球的自转，而是通过推论才知道它的存在。

如果你在一辆火车上向窗外看，你会觉得自己是固定的，乡野的景物却在飞驶而过。而一个站在地里的人却会看到相反的景象：她会觉得自己是不动的，而火车却在运动。当我们说地球围绕太阳转，而不是太阳围绕地球转的时候，其实区别很细微，因为每一种说法都是合乎逻辑的，这要看你到底选谁来做参照物了。如果你把太阳当做参照物，那么地球相对于作为参照物的太阳来说是运动的，而太阳却是不动的。但是，如果把地球作为参照物，就像行星上的生灵所看到的，太阳就是个运动的物体。

如果是这样，那太阳中心说又有什么值得人们大惊小怪的呢？由于在这一问题上与教会的意见相左，可怜的布鲁诺就被教会烧死了。他只是为了一个错误的见解而死的吗？

他的死并不全因支持太阳中心说。布鲁诺提出了很多被教会视为异端的观点——比如上帝不存在之类的问题。就算不支持太阳中心说，他的命运很可能还是一样的。“地球围绕太阳转”取代“太阳围绕地球转”有着重要的意义。以太阳作参照物，让对行星的运动进行数学表述比以地球作为参照物时变得简单了许多。美比单纯的真更有意义。很多观点都可以真实地描述自然，但是总有些观点比别的观点深刻。

如果所有的运动都是相对的，那也就没有什么东西是绝对静止的。牛顿力学都遵从下面这一假设：所有的惯性参照系都是同等有效的。但这对麦克斯韦的方程来说是不正确的。

*

19世纪将近尾声的时候，关于另一个问题的可信性，也很自然地引起了人们的重新思考。如果光是在以太中传播的波，那么，以太很可能是静止的。这样，也必然存在一种（相对于以太）绝对的运动，而不是相对运动。但是，这仍然无法解释为什么麦克斯韦的方程是以另一种参照系为标准的。

这里还要讲到对称。从一个参照系变换到另一个参照系，就是进行空间—时间的对称操作。惯性参照系是关于平移对称的；旋转参照系是关于旋转对称的。我们说牛顿定律都在惯性参照系中，实际就是说这些定律在匀速平移的情况下是对称的。由于某种原因，麦克斯韦的方程不符合这种性质。这仿佛是说，有些惯性坐标系比另外一些更具惯性。如果说某些惯性参照系是特殊的，那它们相对于以太肯定是静止的。

这些问题最后变成了两个，一个是物理学问题，一个是数学问题。物理学问题是，相对于以太的运动是否可以通过实验进行观测？数学问题是，麦克斯韦方程的对称是什么样的？

物理学问题的答案是阿尔伯特·米切尔森（Albert Michelson）发现的。他是一名美国军官，他请了假，和爱德华·莫雷（Edward Morley）一起师从亥姆霍兹研究物理学。他们设计了一台灵敏的设备，能够测出光在不同方向上的传播速度的细微差异，但结论却是没有任何差异。这说明，要么地球相对于以太是静止的（这样，地球围绕太阳转就说不通了），或者根本就没有什么以太，而且光的运动也不遵循一般相对运动的规律。

爱因斯坦从数学的角度研究了这一问题。在他的论文中，并没有

提到米切尔森—莫雷的实验，虽然他后来曾提到了这个实验，而且明言自己的想法受到了它的影响。他不像其他人那样热衷于实验，而去研究麦克斯韦方程的对称，这些方程有一种新奇的特性：它们同时牵涉到空间和时间（爱因斯坦并没有专门研究过对称，但也不是从未涉足）。这些古怪的对称牵涉到无法观察相对于以太（如果真的存在这样一种介质）的匀速运动的问题。

爱因斯坦的理论之所以叫“相对论”，是因为它作出了关于相对运动和电磁的出人意料的预言。

*

“相对论”这个名字很不恰当。这会对人们产生误导，因为爱因斯坦部分理论的最重要特征就是，有些东西并不是相对的。尤其是光速，它是绝对的。如果投射出一束光，站在地面上的人和坐在火车上的人所测到的光速，是一样的。

想要直观地理解这一说法很难，而且乍看起来，简直有点荒诞不经。光的速度是每秒钟 186000 英里，很明显，这就是站在地面上的所测到的光速，但是坐在火车上的那个观测者会得到什么样的结果呢?

我们可以假设火车行驶的速度是每小时 50 英里。首先，想像一下另一辆也以每小时 50 英里的速度与之平行行驶的火车，你可以透过车窗看一下这辆火车的运动速度有多高。

如果这两辆火车是往相同的方向行驶，那么答案就是每小时 0 英里。第二辆火车会和你的火车同步前进，它会一直在你的火车旁边，相对你的火车，它是不动的。如果它以相反的速度行驶，就会以每小

时 100 英里的速度一闪而过，因为它本身的速度还要加上你所坐火车的每小时 50 英里的速度。

如果以这种方法去观测火车的速度，你就会得出这样的结果。

现在我们可以把另一辆火车换成一束光。转换一下光速的计量单位，就是每小时 670616629 英里。如果你的火车自光源向外行驶，你会算出一个相对速度，670616629−50=670616579（英里 / 小时），因为光必须“赶上”火车。另一方面，如果你的火车是朝着光源行驶，你会认为光与火车的相对速度是 670616629+50=670616679（英里 / 小时），因为火车的速度必须与光的速度相加。

根据爱因斯坦的理论，这两个得数都是错误的。这两种观测方式所得到的结果，都还是每小时 670616629 英里，和站在地面上的观测者得到的结果是完全相同的。

这种说法简直不可理喻。如果牛顿的相对运动定律对另一辆火车来说是有效的，那它为什么对光是无效的呢？爱因斯坦的答案是，以极快的速度运动的物体所遵循的物理定律不是牛顿的物理定律。

更确切地说，这些物理定律和牛顿时代的物理定律是不同的。但是这些定律只在物体的运动速度快得接近光速时才会发生作用。在每小时 50 英里这样的低速中，牛顿的定律是爱因斯坦提出的替代定律的一个很好的近似值，以至于根本就看不出二者之间的任何差别。但是随着速度的提高，差别就会变得越来越明显。

这一观点的物理学基础就是麦克斯韦方程的对称，这些对称不但支撑着麦克斯韦的方程本身，还支撑着光速问题。事实上，光速问题就内嵌于方程之中。所以，光速一定是绝对的。

因此，“相对论”的叫法颇具讽刺意味。爱因斯坦实际上想叫它“不变论”（Invariantentheorie）：关于不变的理论。但是“相对论”的

叫法已经被固定下来，并且，在某种意义上说，数学中已经有了所谓的不变理论，所以，尽管爱因斯坦喜欢这个名字，但很可能会被大家弄混。但是，在所有的惯性参照系中，用“不变论”来表述光速的不变造成的迷惑，都远不及用“相对论”造成的迷惑。

*

“相对论”造成了离奇的影响。光速成了一个有限的速度，你可以比光速更快，但你发出的信息不可能比光快。《星球大战》中的超光速状态是不存在的。接近光速的状态下，长度收缩了，时间缓慢地爬行，而且质量无限增大。但是你还应该注意到一个有趣的现象，那就是你的测量工具也会收缩，它会变慢（时间会过得非常慢），或者变重。这就是为什么站在地面上的观察者和火车上的观察者会不受相对速度的影响，测得相同的速度。这也是米切尔森和莫雷无法测出地球与以太的相对运动的原因。

在你移动时，所有事物看起来和你不动时没什么两样。物理定律并不能判定你是运动的还是静止的。它可以告诉你你是否在加速，但如果你是以匀速运动，它无法告诉你你的运动速度。

这仍然让人难以理解，但是试验必须用精确的结果说话。相对论的另一成果是爱因斯坦的著名方程 $E=mc^2$，这个方程把质量和能量联系了起来，直接催生了原子弹，但其在这一过程中的作用常常是被夸大了的。

光对我们来说太熟悉了，以至于我们很少想到过它神秘的特性。它看起来没有重量，照射在每个地方，让我们能看见东西。那么，光是什么呢？电磁波。波是什么呢？某种空间—时间的连续统一体，对

于这种统一体，我们会用“我们不知道”这一有趣的说法来描述。20世纪初，人们以为波赖以传播的介质是以太。爱因斯坦之后，我们对以太有了一点认识：它不存在。波不在任何介质中。

我们会发现，量子力学走得更远。不仅光波不在任何介质中，它还认为：一切都是波。并没有某种介质帮助波传播，而是在波通过空间—时间的构造时引起它的震动，而这种构造本身就是由波构成的。

*

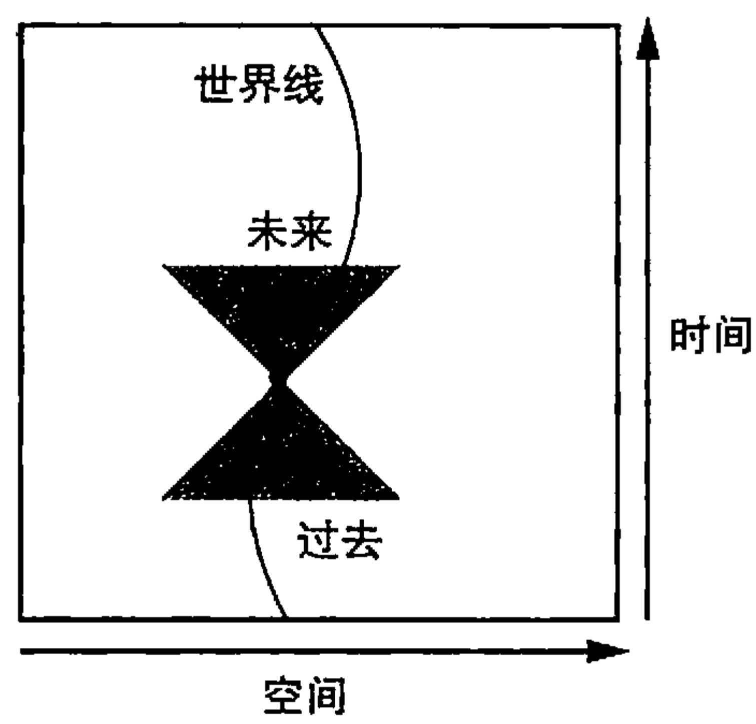
爱因斯坦并不是惟一发现时空的对称不同于典型的牛顿对称的人，比如，这一发现就暗含在麦克斯韦的方程中。以牛顿物理学的观点来看，空间和时间是相互独立又迥然不同的。物理定律的对称是由空间的刚性运动和时间的独立变换组成的。但是就像我前面说过的，这种变换也造成了麦克斯韦方程的可变性。

从这一思考起点出发，数学家亨利·庞加莱和赫尔曼·闵可夫斯基得出了理论数学层面上的一种关于时空对称的新观点。如果他们以物理学的术语表述这种观点，就会驳倒爱因斯坦的相对论，但是他们避开了物理学思考。他们肯定知道，电磁理论的物理定律不会单独地影响时间或空间，而是将对两者的影响混合起来。描述这些相互扭结的变化的数学课题被称为洛伦茨群（Lorentz group），是以物理学家亨德里克·洛伦茨（Hendrik Lorentz）的名字命名的。

闵可夫斯基和庞加莱把洛伦茨群看做对物理定律的某些特性的抽象表达。洛伦茨群也对“时间过得慢了”或“高速运动的物体会收缩”之类与其说是事实不如说是含糊的类推的说法进行了抽象表述。但是爱因斯坦坚持认为这样的说法有着真实的物理学意义。物体和时

间确实具有这样的特点。这样，他构想出了狭义相对论这一物理学原理，把洛伦茨的数学课题变成了对统一的空间—时间（而是不是相互独立的空间和时间）的物理学描述。

闵可夫斯基为这一非牛顿物理学作出了一个几何图形，现在被称为闵可夫斯基时空。这一图形用独立的坐标表示空间和时间，还有一条以随时间变化而移动的微粒画出的曲线（爱因斯坦叫它世界线[world line]）。因为不存在运动速度可以超过光速的微粒，所以世界线在时间的方向上的斜度不会超过45°。这个微粒过去和未来都只能在一个双圆锥范围内，这就是它的光锥。



闵可夫斯基时空的几何图形

这涉及了电和磁这两种基本的自然力，但是没能把另一自然力——重力包括在内。为了提出一种包含重力在内的更普遍的理论，爱因斯坦再次根据自然原理必定对称的原则，提出了广义相对论：时空本身是弯曲的，其曲度和质量有关。这些思想催生了现在流行的宇宙大爆炸理论，认为宇宙是从130亿年前的一个微粒扩张出来的，在这一理论中还包含了著名的黑洞（它质量巨大，连光都无法逃脱它的重力）理论。

*

广义相对论可以回溯到早期的非欧几何研究，在这一研究过程中，高斯得出了“度量”（metric）思想，用以计算任意两点之间的距离。这个公式不是从欧几里得几何学中的毕达哥拉斯定理中来的，就此，诞生了一种新的几何学。只要这个公式遵从一些简单的规则，就会确定一种内涵丰富的“距离”（distance）概念。它遵从的主要规则是，如果 A 点到 C 点中间经过 B 点，A 点到 C 点之间的距离不可缩减。这就是说，从 A 点到 C 点的距离必须等于或小于 A 点到 B 点的距离和 B 点到 C 点之间的距离之和。这就是“三角形公理”（triangle inequality），它之所以叫这个名字，是因为欧几里得几何学认为三角形的任一边的长度都短于其他两条边的长度之和。

毕达哥拉斯公式主导着欧几里得几何，在这种几何中，空间是“平的”（flat）。因此当一种度量不同于欧几里得几何中的度量时，我们就可以将这种区别归因于空间的“弯曲”。你可以想像空间弯曲的样子，但是要想像一个将原空间“弯进去”的更大空间并非易事。思考“弯曲”的更好的办法是，把空间的范围看做是挤扁了的，这样它的内部占有的空间要么比外部占有的空间大，要么就比在外部占有的空间小（BBC“神秘博士”[Doctor Who] 的粉丝可能会想起塔迪斯 [Tardis] 这个宇宙飞船 / 时光机，它的内部比外部更大）。高斯杰出的学生黎曼扩展了这一思想，他使距离得以局部地限定——适用于两个点相距非常近的情况。这种几何被称为黎曼流形，也是最普遍的一种“弯曲空间”。

物理现象不仅仅发生在空间中，而是在空间—时间中，根据爱因斯坦的说法，其中的“平面”几何不是欧几里得式的，而是闵可夫斯

基式的。时间以一种空间的方式进入了“距离”公式。这样的几何结构就是一个“弯曲的时空”。而这正是这位专利局职员所要研究的。

*

爱因斯坦花了很长时间，想给广义相对论构想出一个方程。他先是研究了光在重力场中的运动，结果，这一研究为他另一个独立的基础理论——守恒理论（equivalence theory）奠定了基础。在牛顿力学中，重力有一种能力，使物体之间相互吸引。这种力量产生了加速度。根据守恒原理，加速度总是和重力场的影响密不可分。换句话说，要把重力引入相对论，就必须理解加速度。

1912 年，爱因斯坦说服自己承认，甚至在洛伦茨变换中，重力理论也不可能是对称的。只有当没有物质、重力为零的情况下，这种对称才有可能存在，时空才有可能是闵可夫斯基式的。放弃了对“洛伦茨不变性”的探究，爱因斯坦省去了大量无谓的努力。“我惟一能够确定的，”他于 1950 年写道，“就是在基本方程中，一定会用到守恒理论。”但他同样意识到了守恒理论的局限性：它只能局部地使用，只能无限近似于理论本身。

1907 年，爱因斯坦的朋友格罗斯曼当上了 ETH 的几何学教授，并说服爱因斯坦在那儿谋了一个职位。不久（一年之后），他就离开那里去了柏林，后来又去了布拉格。但是他仍然和格罗斯曼保持着联系，并因此获得了大量的报偿。1912 年，格罗斯曼帮助爱因斯坦找到了他将会用到的数学内容：

我一直无法解决这个问题……我猛然想起高斯关于面的理论

就是揭开这个秘密的钥匙……但是，我当时并不知道黎曼已经以一种更加深入的方式研究了这一几何的基础……当我从布拉格返回苏黎世，我的好友格罗斯曼在那儿，我就问他黎曼的理论是不是可以解决这个问题。

“里奇”就是格雷戈里奥·里奇－库尔巴斯托罗（Gregorio Ricci-Curbastro，1853—1925），他和他的学生图里奥·列维－奇维塔（Tulio Levi-Civita）共同发明了黎曼流形的微积分。里奇张量是曲度的计量方法，比黎曼原来的观念更简单。

据其他资料记载，爱因斯坦曾对格罗斯曼说：“你一定要帮我，不然我就疯了！”格罗斯曼答应了。爱因斯坦后来写道，他“不但帮我省去了相关数学著作的学习，还在重力场方程的研究上给了我很多支持”。1913 年，爱因斯坦和格罗斯曼发表了两个人合作的第一项成果，这篇文章的结尾，是关于他们所寻找的重力场方程的形式的猜想：应力能量张量必须和……（某种东西）成正比。

这种东西到底是什么呢?

他们当时并不知道。它一定是另一种张量，曲度的另一种计量方法。

在这一点上，他们两个都犯了数学错误，这让他们做了很多的无用功。他们相信自己的理论在某种恰当的情况下（平面的时空，极小的重力）符合牛顿的重力理论，这种看法没错。从这一点出发，他们接着推论预想中的方程的规定参数，这种参数很自然地就是他们要找的“某种东西”。但是他们的论证是不合理的，因此所求的规定参数也是无效的。

爱因斯坦相信，正确的重力场方程应该对度量（时空的距离公

式，它决定着自身的所有几何性质）的数学形式起到独一无二的决定作用。这很明显是错误的：坐标系统的改变会使方程发生改变，却不会影响空间内在的曲度。但是爱因斯坦并不清楚所谓的比安基恒等式（Bianchi identity），它阐明了惟一性的不可能，很明显，格罗斯曼对此也不了解。

这是每一位研究者的噩梦：本指望一个看似万无一失的想法会引领自己登堂入室，结果却误入歧途。想要根除这种错误简直比登天还难，因为你开始时认为它并没有错。人们常常意识不到自己已经作出了某种假设。

1914 年底，爱因斯坦终于发现，重力场方程并不是度量的惟一决定者，因为坐标体系是可选择的，坐标体系并没有物理学含义，但却会改变度量公式。他至此仍不知道比安基恒等式，但现在也用不着它了。他最终发现，自己可以完全自由地选择便于使用的坐标系。

1914 年 11 月 18 日，爱因斯坦关于重力场方程的研究又向前推进了一步。他离自己的最终构想已经很近了，并在此基础上作出了两个预测。其中一个解释了水星运行轨道的细微变化，这其实可以说是一个“后见之明”，因为他是在事实发生后作出的解释。行星最靠近太阳的位置（近日点）是缓慢变化着的。爱因斯坦通过自己的重力理论，推算出了近日点的移动速度，他的推算是完全正确的。

第二个预测需要一种新的观测材料来检验其正确性，这一预测是前所未有的，而新的观测材料是检验新理论的最佳途径。根据爱因斯坦的理论，重力会使光弯曲。这种效应的几何原理很简单，并且与测地线（两点之间的最短路径）有关。如果你把一条线拉紧，并将其悬于半空中，它就会形成一条直线。这是因为，在欧几里得空间中，一条直线就是一条测地线。但是，如果你沿着一个足球把这条线拉紧，

它就会形成一条附在足球表面上的曲线。测地线在弯曲空间（球）中本身就是弯曲的。在弯曲时空中也是一样，只是在一些细节上会有些细微的差别。

*

这种效应在自然界中的表现也是很明显的。一颗恒星，比如太阳，会使一些从附近经过的光弯曲。在当时，观察这一效应必须等到日食的时候，这时，太阳的光不会淹没靠近太阳边缘的恒星发出的光。如果爱因斯坦是正确的，那么只要恒星和太阳的位置不重叠，人们观测到的恒星位置就会出现细微的偏移。

这一现象的定量分析却并不是那么一目了然。1911 年，爱因斯坦曾尝试性地预测，恒星位置会发生 1 弧秒的偏移。牛顿也可能会根据自己光是粒子的理论作出类似的预测，重力会吸引粒子，从而使它的运行路径弯曲。但是，1915 年，根据其新的理论，爱因斯坦推算出，光会弯曲 1.74 弧秒，接近原来的预测值的两倍。

现在是在牛顿和爱因斯坦之间作出决断的时刻了。1914 年 11 月 25 日，爱因斯坦写下了自己重力场方程的最终形式。这些“爱因斯坦方程”构成了广义相对论的基础——重力的相对性理论。他们是以一种被称作“张量”的数学形式（一种被众人热捧的矩阵）写成的。爱因斯坦的方程告诉我们，爱因斯坦张量和应力能（Stress-energy）张量的变化率成正比。这也就是说，时空的曲度和物质的量成正比。这些方程遵循着一种对称原则，但只是一种局部的原则。在一个小的时空范围内，只要把曲度的影响考虑在内，它们和狭义相对论的对称就是一样的。

爱因斯坦解释说，他对水星近日点的运动和经过恒星的光的偏折的推算，不会因自己所做的细微的修补而改变。他把自己的方程提交给了普鲁士学院（Prussian Academy），但却发现数学家大卫·希尔伯特（David Hilbert）已经提交过同样的方程，而且不止把他们看成研究重力场的工具。事实上，他认为这些方程还包含了电磁方程，但这种看法并不正确。我们再一次发现，一个顶尖级的数学家差一点打败爱因斯坦，这是个很有意思的现象。

为了验证爱因斯坦关于光线因太阳的重力场而偏折的预测，人们做了好几项努力。第一次是在巴西，却因为下雨而失败了。1914年，一个德国考察队到克里米亚观测日食，却很快就因第一次世界大战而被责令回国。部分成员照做了，可还有一些被逮捕，但最终还是安然无恙地回国了。当然，观测没搞成。这次战争还阻碍了1916年委内瑞拉的观测活动。美国人在1918年又尝试了一次，但没有获得定论性的结果。最终，1919年5月，一支由亚瑟·艾丁顿（Arthur Eddington）率领的英国考察队成功了，但直到11月才发布了他们的观测结果。

这项结果对爱因斯坦的理论比对牛顿的理论更为有利。确实有偏折发生，且偏折程度之大非牛顿的理论所能解释，却与爱因斯坦的理论形成了漂亮的对应。

回过头看，当时的实验并不具备那么大的决定性。实验的误差幅度很大，顶多可以说：爱因斯坦可能是正确的（现在的很多观测，技术和设备都比当时的好，也证实了爱因斯坦的理论）。但在当时，这被看成了具有决定意义的事件，被媒体发了疯似的报道着。谁要能证明牛顿是错误的，他就一定是个天才；谁要是发现了一种全新的物理学，就一定是最伟大的在世科学家。

一个传奇就此诞生。爱因斯坦在伦敦的《泰晤士报》（*Times*）上阐述了自己的思想。几天后，报纸的评论版这样写道：

> 对于我们信心满满的安稳感来说，它是一种崭新而强烈的震动和忧虑，甚至乘法表都不再安稳……这将要求我们两个皇家协会的主席对光有重量、空间是有限的的观点作出有说服力，哪怕是勉强可以接受的解释。它不仅是上述观点的一种反对意见，还是这些观点的终结。尽管它可能是给更高级的数学家看的，但它也是广大民众的理论。

事实证明，它是给更高级的数学家们看的。不久，《泰晤士报》就说，世界上"只有 12 个人可以理解这位'一夜成名的爱因斯坦博士'的理论"。这一神话流传了很多年，甚至在很多物理学研究生的课程中都是这样讲的。

1920 年，格罗斯曼第一次出现了多发性硬化症状。1930 年，他写出了自己平生最后一篇论文，并于 1936 年去世。爱因斯坦却渐渐变成了 20 世纪的标志性科学家。到了晚年，爱因斯坦才看淡了自己的名声，依稀觉得这东西有点可笑。但是此前，他似乎很乐于跟媒体打交道。

但我们对爱因斯坦的介绍必须到此为止，至多再谈一下那项他在 1920 年后所做的毫无建树的研究——用"统一场论"（unified field theory）融合相对论和量子力学。他一直在进行这项研究，直到 1955 年去世。

量子论的五重奏

“所有的发现都已经完成，只剩下一些补遗的工作可做了。”对于一个立志研究物理学的青年才俊来说，这真是个令人丧气的消息，尤其是当这句话出自一个物理学教授之口时。而我们应该记住这个教授的名字，他就是菲利普·冯·约利（Philipp von Jolly，1809—1884）。

当时是1847年，冯·约利的观点反映了那时大多数物理学家的看法：物理学已经完成了。1900年，大名鼎鼎的开尔文说：“物理学中已经没什么可发现的了，只剩下越来越精密的测量了。”

他还说了这样的话：“我可以很伤心地肯定，任何比空气重的空中飞行器都是不可能的”，并且，“人类要登录月球还有太多难题需要应付，而解决这些问题还要花200年的时间。”开尔文前半段的研究是正确的，而后半段的研究却是错误的。

但是，他没有全错。在他1900年的演讲“19世纪留下来的关于热和光的动力理论的疑问”中，开尔文阐述了当时物理学中的两大关键性漏洞：“漂亮而简洁的动力理论认为热和光是两种运动模式，但

这一理论现在被两个疑问弄得模糊难解。第一个疑点是，地球是如何在一种弹性体（本质上说，即以太）中运动的？第二个是麦克斯韦－玻耳兹曼原理对能的划分。”第一个疑问催生了相对论，第二个则催生了量子理论。

幸运的是，约利观点的年轻接受者并没有因此被吓倒。他没有想过要什么新事物，他说，我只是想对基础物理学了解得更透彻一些而已。而就在这一研究过程中，他作出了20世纪物理学中两个最重大的革命性发现之一，扫除了开尔文的第二个疑问。他就是马克斯·普朗克。

*

尤利乌斯·威尔海姆·普朗克（Julius Wilhelm Planck）是基尔大学（Kiel）和慕尼黑大学的法学教授。他的父母都是神学教授，兄弟是一名法官。所以，当第二任妻子埃玛·帕齐格（Emma Patzig）生下一个儿子时（这是他的第六个孩子了），这个孩子就注定要在知识分子的环境中长大。马克斯·卡尔·恩内斯特·路德维希·普朗克（Max Karl Ernest Ludwig Plank）生于1858年4月23日。这时的欧洲政局动荡，在这个男孩的记忆里，还有1864年丹麦与普鲁士的战争期间，普鲁士和奥地利军队开进基尔的情景。

1867年，普朗克一家搬到了慕尼黑，马克斯也进入了麦斯米兰国王学校（King Maxmilian School），跟着赫尔曼·穆勒（Hermann Müller）学习数学。穆勒教他天文学、力学、数学，还有一些基础的物理学，包括能量守恒理论。普朗克是一个聪明的学生，在16岁时就提前毕业了。

他还是个出色的音乐家，但是，虽然约利已经讲过自己对物理学的观点，他还是选择了物理学。普朗克在约利的指导下完成了几项实

验，但很快就转向了理论物理学。他和世界顶尖级的物理学家和数学家们过从甚密，又在1877年去了柏林，跟着亥姆霍兹、古斯塔夫·基尔霍夫（Gustav Kirchhoff）以及魏尔斯特拉斯学习。1878年，他通过了第一次的考试，1879年，又以一篇热力学论文获得了博士学位。有一段时间，他在一所老旧的学校里教授数学和物理学。1880年，他关于物体在不同温度下的均衡状态的论文受到了认可，从而得以开始了自己辉煌的学院生涯。他正式获得了一个教职，但直到1885年，他才成了基尔大学的副教授。他的研究集中在热力学，尤其是熵理论上。

普朗克认识了一个朋友的妹妹玛丽·梅尔克（Marie Merck），1887年，两个人在一间租住的公寓里结了婚。他们一共生了四个孩子：卡尔、双胞胎爱玛和格丽特，以及埃尔文。

1889年，也就是那对双胞胎出生的那一年，马克斯接替了基尔霍夫在柏林的教职，并于1892年成了正教授。他们举家搬进了柏林古耐沃德地区（Grunewald）的一栋别墅，和另外几位顶尖级的学者比邻而居。其中一位就是哈纳克（Adolf von Harnack），他们成了好友。普朗克很爱交际，很多著名的知识分子经常到他家里造访。其中就包括爱因斯坦、奥特·哈恩（Otte Hahn）和利塞·迈特纳（Lise Meitner），迈特纳后来作出了核裂变的几项根本性发现，其中一部分经过发展，成了原子弹的理论基础。在这些社交活动中，普朗克继续着演奏音乐的传统，这一传统是亥姆霍兹开创的。

普朗克的生活一度非常美满，但后来玛丽罹患肺病，很可能是结核，并于1909年去世。一年半后，52岁的马克斯又结婚了。这次是和玛德·冯·赫斯林（Marga von Hoesslin），两个人又生了他的第三个儿子赫尔曼。

*

1894 年，地方上的一家电气公司想研发一种更高效的灯泡，因此，马克斯开始了一些合同性的工业研究。理论上讲，灯泡分析是物理学中典型的“黑体辐射”（blackbody radiation）问题—— 一个完全非反射性的物体如何发光。这样的物体被加热以后，会发出全频率的光，其强度和能量都会随着频率变化。根本问题是，光的频率是如何影响其强度的？少了这些基本数据，要发明一种新的灯泡就太难了。

古典物理学基本原理中有很好的实验结果，以及一条理论定律，也就是瑞利－金斯定律（Rayleigh-Jeans law）。可惜这条定律和高频率情况下的实验并不相符。事实上，这预言了某种不可能性：随着光的频率的升高，它的能会变得无限大（所以实验不可能成功）。这种不可能性的结果后来被称作“紫外灾变”。而进一步的实验导出了一条新定律，适用于观测高频率辐射。这条定律的发现者是威廉·维恩（Wilhelm Wien），因此被命名为维恩定律。

但是，对于低频率辐射来说，维恩定律是错误的。

物理学家们面前摆着两条定律：一个适用于低频辐射却不适合高频辐射，另一个则正好相反。普朗克决定调和这两个定律，他写下了模拟低频辐射的瑞利－金斯定律和高频辐射的维恩定律的数学表达式。这个最终的公式现在被称为黑体辐射的普朗克定律。

这个巧妙的新定律和实验配合得非常严密，概括了全部的电磁辐射光谱。但是这条定律纯粹是经验主义的，因为它是从实验中得来的，而不是来自物理学原理。普朗克原本只是想多了解一些物理学，这个目的已经实现，但为了找到新公式的物理学原理，他仍然继续着自己的研究。

最终，在1900年，普朗克发现了这个方程的一个奇特性质。只需对瑞利－金斯定律作出一些细微调整，就可以用相同的推算方法推导出这个方程。古典物理学的推导方法假设，原则上来讲，任何给定频率的电磁辐射的能都可以获得一个值。在特殊情况下，你可以把它设定为一个近似于零的值。普朗克发现，这个假设就是造成紫外灾变的原因，而如果自己作出了另一种假设的话，这个令人头疼的无限性就消失了。

尽管说这只能是一个假设，但也必须具有根本性。一个给定频率的辐射的能必须是整数个固定大小的“捆”（packet）。事实上，每一捆的大小都必须和频率成正比，也就是说，必须等于频率和某个常数的乘积。这个常数就是我们常说的普朗克常数，其表示符号是 h。

这些能量捆叫做“量子”（quanta，单称的 quantum）。普朗克把光量子化了。

那么，为什么试验者从来没有发现过能总是整数个量子呢？通过对比自己的推算和试验中观测到的能，普朗克推算出了这个常数的大小，这是一个非常非常小的数。事实上，h 大约相当于 6×10^{-34} 焦／秒。大致来说，要发现可能存在的能量“误差”的幅度（古典物理学允许这种程度的误差，而量子物理学中却决不允许），你的观察数据必须精确到小数点后34位。即使在今天，也很少有可以精确到超过小数点后六位或七位的物理量，而在当时，能精确到小数点后三位已经很奢侈了。要对量子进行直接的观测，需要达到的精确度是难以想像的。

这么一个细微得难以觉察的数学误差竟可以对辐射定律产生这么大的影响，看起来简直有些不可思议。但是，公式的推算还必须把所有可能的频率对能的总体影响考虑在内，其最后结果其实就是所有量

子的综合影响。站在月球上，你看不见任何一粒单个的沙子，但却可以看见撒哈拉沙漠。足够多的小单位组合起来，其结果可能相当巨大。

普朗克的物理学事业兴旺发达，个人生活却充满了悲剧。他的儿子卡尔在第一次世界大战中战死；女儿格丽特 1917 年死于难产；爱玛嫁给了格丽特的鳏夫，1919 年也遭遇了同样的命运。晚些年，埃尔文因在 1944 年参与了刺杀希特勒的行动而被纳粹处决。

*

1905 年，支持普朗克看法的新证据出现了，这个证据就是爱因斯坦的光电效应著作。大家应该记得，光可以转换成电是一个新发现。爱因斯坦很清楚，电是分立封装的（discrete package）。实际上，人们正是自那时起开始知道，电是一种名为“电子”的微粒的运动。通过对光电效应的研究，爱因斯坦推论，光也一定存在这种情况。这不但验证了普朗克的光量子思想，还解释了光量子到底是什么——和电子一样，光波也必定是粒子。

波怎么会是粒子呢？但这确实是实验得出的结果。光粒子（光子）的发现，很快就催生了一个世界的量子绘景（quantum picture），在这个世界中，粒子就是波，有时是波，有时是粒子。

量子越来越受到物理学界的重视。伟大的丹麦物理学家尼尔斯·波尔（Niels Bohr）提出了原子的量子化模式，在这种模式中，电子沿圆形轨道绕着原子核运动，一定大小的轨道容纳着一定量的量子。法国物理学家德布罗意（Louis Victor due de Broglie）推想，如果光子既是波又是粒子，又因为合适的金属受到光子的影响就会放射

电子，那么，电子也一样既是波又是粒子。事实上，所有的物质都是双重的存在，有时是密实的粒子，有时是起伏的波。

无论是“粒子”还是“波”，都没能恰当地描述极小的物质。物质的终极构成是二者的混合体——波粒子。德布罗意还发明了描述波粒子的公式。

现在到了关乎本书主题的关键部分。埃尔温·薛定谔（Erwin Schrodinger）把德布罗意的公式变成了描述波粒子运动的方程。和牛顿的定律是古典物理学的基础一样，薛定谔的方程成了量子力学的基础。

*

埃尔温1886年生于维也纳，是个混血儿。他的父亲鲁道夫·薛定谔的工作是加工裹尸布，同时也是个植物学家。鲁道夫是个天主教徒，而埃尔温的母亲乔金·埃米利亚·布琳达（Georgine Emilia Brenda）却是路德派教徒。1906年至1910年，埃尔温在维也纳学习物理学，师从弗兰茨·瑟拉芬·埃克斯纳（Franz Serafin Exner）、弗雷德里希·哈瑟诺尔（Friedrich Hasenöhrl），并于1911年成了埃克斯纳的助手。正是这一年，第一次世界大战爆发，他加入了奥地利的炮兵部队。第一次世界大战结束后，他和安妮玛丽·伯蒂尔（Annemarie Bertel）结了婚。1920年，他成了斯图加特大学的副教授，又在1921年成了布雷斯劳（Breslau，今天的弗罗茨瓦夫，属于波兰）大学的正教授。

他在1926年发表于一篇论文中的方程现在是以他的名字命名的，在这篇论文中，他写出了氢弹的辐射频谱的能量级。紧接着，他又写

下了三篇重要的量子物理学的论文。1927 年，他成了普朗克的同事，但是在 1933 年，纳粹的反犹主义让他十分沮丧，他离开德国去了牛津大学，成了玛德林学院（Magdalen College）的教员。到那儿不久，他就和保罗·狄拉克（Paul Dirac）共同获得了诺贝尔物理学奖。

薛定谔一直过着绯闻不断的反传统生活，他和两个女人一起生活，这触怒了那些温文尔雅的牛津人。他在一年中搬了两次家，最后去了普林斯顿大学。普林斯顿大学给了他一个显赫的职位，可他却拒绝接受，这可能是因为在这儿他不能和两个妻子及同在家中的情妇保持紧密的联系，而在牛津就好得多。最终，1936 年，他不顾奥地利的刻板风气，定居到了格拉茨（Graz）。

希特勒占领奥地利给薛定谔这个纳粹的对头带来了很多麻烦。他只好公开放弃了自己之前的立场（很久以后他曾为此向爱因斯坦道歉）。他的骑墙作风并未奏效，最终还是由于政治不坚定而丢掉了工作，逃亡到了意大利。

薛定谔最终定居在都柏林。1944 年，他出版了《生命是什么?》（*What Is Life?*）一书，试图用量子物理学解释生命现象，但是漏洞百出。这本书的思想建立在与热力学第二定律相反的“负熵”概念的基础上。薛定谔强调，生命体的基因一定是某种复杂的分子，内含编码指令。我们现在称这种分子为 DNA，但是 DNA 的结构是 1953 年由弗朗西斯·克里克（Francis Crick）和詹姆斯·华生（James Watson）发现的。他们在一定程度上受到了薛定谔的启发。

在爱尔兰，薛定谔对性的态度仍然很开通，他爱上了自己的学生，还和两个女人各生了一个孩子。他于 1961 年在维也纳去世，死于肺结核。

*

薛定谔的猫也很出名。这并不是真实的猫，而是一个思想实验。这常被看做人们不考虑薛定谔的波是否真的是物理事物的原因。相反，它们被看成一种具有相应的因果效应，但永远无法用实验证实的后台描述。但是，这种说法是自相矛盾的，如果这些波不存在，那么，它们的效应为什么总会这么顺利地得到实现呢？

撇开这个问题不谈，我们再回头看看薛定谔的猫。根据量子力学，波粒子会相互影响，尖端相遇时，它们会得到强化，一方的尖端和一方的侧面相遇时，会相互抵消。这种作用叫做“叠加”(superposition)，所以量子波粒子可以相互叠加，也就是说它们可能存在一种潜在状态，在这种状态下，它们既非此又非彼。事实上，根据波尔对量子理论的理解和著名的“哥本哈根解释”，这就是事物的本征态（natural state)。只有当我们观测一个物理量时，才会使量子脱离叠加状态，进入一种“单纯”状态。

这种解释对于电子来说是完全有效的，但薛定谔不知道对一只猫会不会有效。在他的思想实验中，一只锁在箱子里的猫就处于生和死的叠加状态。当你打开箱子观察时，才能决定猫的生死。特里·普拉切特（Terry Pratchett）在《假面舞会》(*Maskerade*）的注释中说，这只猫不会处于叠加状态。一只名叫“格里波”的猫在箱子里出现了第三种状态——呜呜哀号。

薛定谔也知道猫不会出现叠加状态，但却不是通过这样的方式知道的。电子是亚微粒，它的存在方式是量子式的。如果我们对它进行测量，它会拥有一个特定的位置、体积或一种旋转。猫是一种可见物，其存在方式不是量子式的。电子的状态可以叠加，但猫的状态就

不行。我和妻子有两只猫，但当它们想要相互叠加时，就变成了两只发怒的猫，还有它们四处飞散的皮毛。在量子理论的术语中，这叫“脱散”(decoherence)，指的是猫之类的较大的量子系统，这种量子系统就是我们日常生活中熟悉的“经典”系统。脱散的概念告诉我们，猫包含的波粒子太多，以至于这些波粒子一转眼的工夫就纠缠在一起，比光通过电子的直径的时间还要短，因此不可能发生重叠。由于猫是一种由巨量的量子微粒组成的可见系统，所以不可能出现重叠现象。它们可以是活的，也可以是死的，但绝不可能既是活的又是死的。

但是，如果一个东西足够小（我们所说的东西非常的小，并不是那种通过普通的显微镜就可以看到的东西），其运转形式就会符合量子理论，可以在同一时间内处于不同的状态。这种理论改变了一切。

*

奇异的量子世界是通过维尔纳·海森堡（Werner Heisenberg）的研究才呈现在世人面前的。海森堡是个杰出的理论物理学家，但是却对实验很不在行，以至于在博士学位考试时，竟回答不了望远镜和显微镜之类的简单问题。他甚至连电池组的工作原理都不知道。

奥古斯特·海森堡（August Heisenberg）与安娜·魏克林（Anna Wecklein）于 1899 年成婚。奥古斯特是一个路德会信徒，而魏克林是一个天主教徒，但是为了能顺利结婚，她也改宗为路德会信徒。两个人有着很多的共同点：奥古斯特是一名教师，同时是古希腊方面的古典学专家；魏克林是一名校长的女儿，也是古希腊悲剧的专家。他们的大儿子厄尔文出生于 1900 年，成了一名化学家；而生于 1901 年的

二儿子维尔纳，则改变了世界。

当时的德国仍然是一个君主制的国家，教育这个职业的社会地位很高，所以海森堡夫妇生活优裕，可以把自己的儿子送到优秀的学校去读书。1910年，奥古斯特成了慕尼黑大学的教授，研究中世纪和现代希腊。1911年，维尔纳进入了慕尼黑的麦克米兰中学，普朗克也曾在那里读书。维尔纳的外祖父尼古拉斯·魏克林（Nikolaus Wecklein）就是这所学校的校长。维尔纳聪明伶俐，显示出了过人的数学和科学才能，这一定程度上是由于他的父亲一直鼓励他向自己的哥哥看齐。他还有很好的音乐天赋，钢琴弹得非常好，12岁时就能在学校的音乐会上表演了。

海森堡后来写道，“语言和数学很早就唤起了我的兴趣，”他的拉丁语和希腊语都是优等，数学、物理和宗教学成绩也都很好，最差的科目是体育和德语。他的数学老师克里斯托弗·乌尔夫（Christoph Wolff）是个了不起的人物，他一直通过专门给海森堡出题来督促他的学习。很快，这个学生就超越了自己的老师，海森堡的学校报告这样写道：“他在数学—物理学方面的独立研究已经远远超过了学校的要求。”他自学了相对论，更喜欢其中的数学内容，而不是其中的物理学成分。父母让他去辅导一个地方学院的学生，以帮助她通过考试，可是维尔纳却自作主张地给她讲微积分，跟学校的考卷没有一点关系。他还对数论产生了兴趣，“（数论）如此清晰，所有的东西都让人一目了然。”

为了提高维尔纳的拉丁语水平，父亲给了他一些用拉丁语写成的古老的数学著作，其中就包括克罗内克的一篇关于数论的论文。克罗内克是一个世界级的数论家，他有一个信条非常著名：“上帝只创造了整数，其余都是人类的发明。”海森堡受到这篇论文的激发，尝试

着去证明费马最后定理。九年后，海森堡以优异的成绩从中学毕业，进入了慕尼黑大学。

第一次世界大战爆发后，协约国封锁了德国。食物和燃料非常短缺，由于无法供暖，学校被迫关闭。维尔纳一度饿得极度虚弱，甚至从自己的自行车上掉到了路边的壕沟里。他的父亲和他的老师坚持和军队作斗争；而年轻人想要存活就不得不接受国家主义教条。第一次世界大战的结束也是德国君主制的末日，巴伐利亚的社会主义者们很快建立了苏维埃政府。1919 年，柏林来的德国军队赶走了社会主义者们，但保留了一定程度的社会民主。

由于德国的战败，维尔纳和很多同代人一样，抱怨长辈们的军事失败。他成了一个团体的领袖，这个团体和妄图恢复君主制、梦想建立第三帝国的极端右翼组织“新童子军”（New Boy Scout）联系密切。“新童子军”的很多分支都是反犹主义的，而维尔纳的团体中却有一部分犹太人。他花了很多时间，和这些男孩一起宿营、徒步旅行，还常常一起通过浪漫的想像重温德国曾经的辉煌。但是，1933 年希特勒上台后，这些活动都停止了，因为除了自己建立的组织，他禁绝了所有青年组织。

1920 年，维尔纳去了慕尼黑大学，打算当一名理论数学家，但是后来，在与一位理论数学教授会面之后，他就打消了这个念头。他决定转向物理学研究，师从阿诺·索末菲尔德（Arnold Johannes Wilhelm Sommerfeld）。索末菲尔德很快就发现了维尔纳的才能，允许他加入了高级班。不久，维尔纳就完成了一些很有独创性的、关于原子结构的量子物理学研究。他 1923 年就获得了博士学位，打破了学校的纪录。也是在同一年，希特勒策划了“啤酒店暴动”，企图推翻巴伐利亚政府，并打算以此作为进军柏林的前奏，但是，这次暴动失败了。

接着，通货膨胀飞速蔓延，德国濒临崩溃。

维尔纳继续着自己的物理学研究。他和很多顶尖级物理学家合作研究量子理论，因为在当时，量子理论是个热门话题。维尔纳和马克斯·伯恩（Max Born）试图发明一种更高级的原子理论。海森堡忽然想到，可以用原子的光谱频率来描述原子的状态。他将这一思想总结到了一个特殊的数字串中。伯恩发现这种数字串非常有用，数学家们把它叫做矩阵。伯恩觉得这种思想很有意义，于是就把它拿去发表。随着思考的深入，他们将其发展成量子理论的一种新的对称数学：矩阵力学。矩阵力学常被看做是薛定谔的波动力学的竞争理论。

*

到底哪种理论是正确的呢？事实上，两种理论都得到了验证，1926 年，薛定谔发现了这一点。它们是对同一种潜在观念的两种不同的数学描述，这和欧几里得作图法与代数学是观察几何的两种殊途同归的方式一样。起初，海森堡对此不以为然，因为他的矩阵法的本质是由电子状态的变化所引起的不连续转移。要了解其矩阵，必须研究相应的能的变化。他无法想像波这种连续的存在物也会出现不连续的情况。在一封写给奥地利—瑞士物理学家沃尔夫冈·泡利（Wolfgang Pauli）的信中，他写道："我对薛定谔的物理学理论想得越多，越是觉得它令人生厌……他说自己的理论的可想见性（visualizability）'也许不太正确'，其实换句话说，简直就是一坨屎。"这种不和其实是另一场更古老的争论的重演，伯努利和欧拉曾经在波方程的解法问题上发生过争执。伯努利发现了一种求解公式，而欧拉却无法理解这个貌似连续的公式怎么可能是一种不连续的解法。然而，伯努利是正确

的，薛定谔也是正确的。他的方程可能是连续的，但是这些方程的很多性质都可以是不连续的，其中就包括能量级。

大多数物理学家都更喜欢波动力学的表达方式，因为它更加直观，而矩阵却太抽象了。海森堡却仍然更喜欢自己的数字串，因为它们是由可观测的量组成的，而且，薛定谔的波看起来是不可能通过实验来侦测的。事实上，量子理论的“哥本哈根解释”（可以戏剧化为薛定谔的猫）认为，所有对波的侦测都会使波“分裂”成单一状态，对此界定得已经很清楚了。所以，海森堡越来越关心量子世界的哪些方面是可测量的，如何测量。在海森堡的数字串中，你可以从各个方面入手进行测量，但是，你无论如何都无法对薛定谔的波进行测量。海森堡认为这就是自己坚持矩阵法的重要原因。

沿着这条思路，他发现，原则上讲，对波粒子位置的测量可以要多精确有多精确。但是这种精确是有代价的，因为你对粒子的位置测量得越精确，对其冲量的了解就越不精确。反过来说，当你精确地了解了粒子的冲量，就测不出粒子的位置了。在能量和时间之间，也存在这种两难，你可以测准其中的一个，但不能同时测准二者，除非你压根就没想测太准。

这并不是个实验问题，而是量子理论固有的性质。1927 年 2 月，他把自己的理论写信告诉了泡利。后来，海森堡以这封信为基础写了一篇论文，而他的理论也从此得名——“测不准原理”（uncertainty principle）。这是物理固有的有限性的最重要的例子之一，爱因斯坦所说的没有任何东西的速度可以超过光速则是另一个例子。

1927 年，海森堡成了德国最年轻的教授，任教于莱比锡大学。1933 年，希特勒掌权，海森堡获得了诺贝尔物理学奖。这使他声名大振。由于在纳粹专政期间，他自愿留在德国，导致很多人认为他本人

就是个纳粹党徒。但是可以确证，他并不是纳粹。但他是爱国者，这把他和纳粹联系了起来，还参与了纳粹的很多活动。证据表明，海森堡曾试图阻止把犹太人赶出大学的法令，但是徒劳无功。1937年，他被称为“白人犹太”（white jew），而且有被送到集中营的危险，但是一年后，党卫军的头头海因里希·希姆莱为他洗脱了罪名。还是在1937年，海森堡和一位经济学家的女儿，伊丽莎白·舒梅切尔（Elisabeth Schumacher）结了婚。他们的第一胎是双胞胎，最后一共生了七个孩子。

第二次世界大战期间，海森堡成了德国研究核武器的科学家带头人之一。他在柏林的核反应堆工作，妻子和孩子则被送到了自己家族位于巴伐利亚的夏季住宅中。他在德国原子弹工程中的角色非常矛盾。第二次世界大战结束后，他被英国人关押在剑桥附近的一所房子里，审讯了六个月。他的审讯记录最近得以发表，加剧了世人的争论。海森堡确实说过自己只想制造一个核反应堆（引擎），却不想用它来制造炸弹。“我确实认为我们可以制造一个铀引擎，但是从没想过要制造炸弹，我发自内心地希望它只是一个引擎，而不是一个炸弹。我承认这一点。”对于这种说法的真实性，人们仍然争讼不休。

第二次世界大战后，他在英国的监禁被解除，继续回到了量子理论的研究上，并于1976年死于癌症。

*

在德国，大多数量子理论的创始人都出身知识分子家庭，他们的父亲往往是医生、律师、学者，或从事其他诸如此类的职业。他们生

活在豪华的房子里、演奏音乐，并参与当地的政治和文化生活。而量子理论伟大的英国创始人的童年则悲惨得多。他和自己专横而乖戾的父亲住在一起，而这位父亲和自己的父母也分居两地；这孩子的母亲整天战战兢兢，在丈夫和他们的小儿子在餐厅里默不作声地用餐时，她只能带着两个孩子在厨房里吃饭。

这位父亲名叫查理·阿德里恩·拉迪斯拉斯·狄拉克（Charles Adrien Ladislas Dirac），1866 年出生于瑞士的瓦莱州（canton of Valais），20 岁时就远离家门。查理在 1890 年到了布里斯托，但直到 1919 年才成为英国公民。1899 年，他与一名船长的女儿弗洛伦斯·汉娜·霍尔顿（Florence Hannah Holten）结婚。第二年，他们的第一个孩子雷吉纳（Reginald）出生，两年后，第二个儿子保罗·阿德里恩·莫里斯（Paul Adrien Maurice）出生；又过了四年，他们生了一个女儿，取名贝阿特丽斯。

查理并没有告知自己的父母自己已经结婚，也没有告诉他们其实他们已经当上了祖父母，直到 1905 年他回瑞士探母时，他们才知道上述情况。

查理在布里斯托的“商业技术学院”（Merchant Venturers Technical College）当老师。人们大多认为他是一个好老师，但却出了名地没有人情味，而且非常严厉。总之，他是个严守纪律的人，很多老师也都是这样。

保罗生性内向，而父亲的孤僻与缺乏社会活动更加剧了他的内向。查理要求保罗与自己说话的时候只能讲法语，这可能是出于督促他学习这门外语的目的。由于保罗的法语很糟糕，以至于干脆就不和父亲说话了。他转而把时间都花在了对大自然的遐想上。狄拉克一家之所以缺乏社会活动，也可能是只用法语交谈的习惯所致。人们从来

都不清楚保罗到底是恨他的父亲，还是只不过有点讨厌他，但是在查理死时，保罗的第一个念头是："现在我自由多了。"

查理对保罗的高智商感到很骄傲，对自己几个孩子的期望也很高，希望孩子能够照自己为他们拟定的计划去走。当雷吉纳说自己想当医生的时候，查理坚持说他应该去当工程师。1919 年，雷吉纳获得了一个很不怎么样的工程师学位，五年后，在伍尔弗汉普顿大学（Wolverhampton）大学的一个工程项目工作时，他自杀了。

保罗和父母住在一起，也在雷吉纳那所学校学习工程学。他最喜欢的是数学，却没有选择这个学科。这可能是由于他不想违背父亲的意愿，而且他一直有一种错误的观念，认为学数学将来只能去教书，这种观念现在还很流行。没有人告诉他其实学数学还有很多职业可以选择，比如去做研究。

拯救了保罗的，是一则报纸头条。1919年11月7日的《泰晤士报》的头版赫然地写着："科学的革命，新的宇宙理论，牛顿的思想被推翻"；第二版的中间是文章的副标题："空间'弯曲'"。一时间，所有人都开始谈起了相对论。

我们可以回想一下前面的内容。广义相对论的一个预测就是重力可以使光弯曲，弯曲程度是牛顿定律所能预测出的弯曲程度的两倍。弗兰克·迪森（Frank Dyson）和艾丁顿还为此组织了一个探险队，到西非的普林西比岛（Principe Island ）观察了日食。格林尼治天文台的安德鲁·克罗姆林（Andrew Crommelin）组织了的第二次类似的探险，他们去了巴西的索布拉（Sobral）。两支探险队都对日全食时靠近太阳边缘的恒星进行了观测，发现恒星显现的位置和它本来的位置有一定程度的偏移，其偏移程度和爱因斯坦的预言相符，却不符合牛顿力学。

一夜成名的爱因斯坦给母亲寄了一张明信片，他在上面写道："亲爱的妈妈，我今天很高兴，H. A. 洛伦茨给我发电报说，英国的一个探险队已经证明了太阳光的偏折。"狄拉克被这件事深深吸引了。"相对论让我万分激动。我们谈了很多关于相对论的事情。学生们都在讨论相对论，但却没有什么后续的消息出现。"公众对相对论的理解只是字面意义上的。哲学家们说他们多年以前就知道"所有事物都是相对的"，说这种所谓的新物理学其实早就过时了。可惜，这只能显示他们的无知，以及对不同的术语体系的混淆。

保罗听了几次布里斯托大学的哲学教授查理·布罗德（Charlie Broad）关于相对论的讲座，但是这些讲座的数学内容空洞无物。最终，他买了一本艾丁顿的《空间、时间与重力》(*Space*，*Time and Gravitation*)，开始自学必要的数学和物理学知识。离开布里斯托大学之前，他已经全面理解了狭义相对论和广义相对论。

*

保罗很擅长理论性思考，操作实验却很糟糕。在他晚年的时候，科学家们经常谈到"狄拉克效应"(Dirac effect)：只要他走进实验室，就会犯下大错。如果他当了工程师，那可真是个灾难。由于战后的经济萧条，他发现虽然自己拥有一流的学位，却找不到工作。幸运的是，他得到了到布里斯托大学学习数学的机会，学费全免，这让他欣喜若狂。在布里斯托大学，他专门研究应用数学。

1923 年，保罗成了剑桥大学的硕士研究生，在那儿，腼腆的性格成了他的一个重大障碍。他不喜欢运动，朋友也很少，而且从来不和女性打交道，所以，他把时间都花在了图书馆里。1920 年夏天，他和

哥哥雷吉纳在同一家工厂上班。他们经常步行走过大街，但是从来没有停下来与人交谈过，少言寡语的家庭习惯竟然根深蒂固到了如此的程度。

保罗很快就脱颖而出，还不到六个月，他就写出了第一篇研究论文，另外几篇论文也很快奔涌而至。1925 年，他开始着手研究量子力学。这年秋天，他有一次在剑桥郡的乡村散步，忽然想起了海森堡的“数字串”。这就是矩阵，而且是不可交换的矩阵，有一种东西自始至终困扰着海森堡。李有一种思想，认为在这种情况下，重要的量是换位：$AB-BA$，而不是乘积 AB，狄拉克对李的这一思想非常清楚。同时他的脑海中还一直萦绕着另一种思想，他想起在汉密尔顿的力学公式中也有着类似的观念，这种观念被称为泊松括号。但是狄拉克却没能记起那个公式。

这一夜他一直在思考，几乎就没有睡，第二天早上，“赶紧跑到一个刚刚开门的图书馆，到惠特克（Whitaker）的《分析动力学》（*Analytical Dynamics*）中查找泊松括号，发现那正是我想要的。”他的发现是这样的：两个量子矩阵的换位子和古典物理学中相应的、可变的（通过与常数 $ih/2\pi$ 相乘）泊松括号是等同的。这里的 h 就是普朗克常数，i 就是 $\sqrt{-1}$，π 就是圆周率。

这是一个重大的发现。它告诉物理学家怎样把古典物理学转换成量子物理学。这种优美的数学，把两种深奥却毫无关联的理论联系了起来。海森堡为此大受震动。

狄拉克对量子物理学的贡献很多，我只讲一下他最重要的发现好了。这就是他在 1927 年发现的电子相对理论。物理学家们已经知道了电子会“旋转”（spin），这种旋转和球体绕轴线旋转类似，但也有很重要的不同点。如果你把一个球旋转 360°，把坐标也转动 360°，那

么这个球就又回到了原来的位置上。但是当你把一个电子旋转 360°时，它却是颠倒的。你必须把它旋转 720°，它才会回复原状。

这和四元数非常相似，用四元数解释空间旋转时，也会出现同样的情形。从数学思想上讲，空间的旋转对应的是 SO（3）群，而与四元数和电子相关的群却是 SU（2）。这些群都是差不多一样的，但是 SU（2）比 SO（3）大一倍，在某种特定意义上，它就是两个 SO（3）。这种现象叫做“双覆盖”（double cover），其结果是，把一个 360°旋转扩展成两个 360°旋转。

狄拉克并没有用四元数，也没有用到群，但是在 1927 年圣诞节过后，狄拉克提出了与此二者具有同样作用的“旋转矩阵”。后来，数学家将狄拉克的矩阵普遍化为“旋量”（spinors），这一概念在李群的表示论中非常重要。

通过旋转矩阵，狄拉克构造了电子的相对量子模型。从这个模型中，狄拉克获得了自己希望得到的一切，甚至还超出了狄拉克的预期。它预示了负能量和与之对应的正能量理论的提出。在经过几次碰壁之后，狄拉克提出了“反物质”（antimatter）概念——所有的粒子都有相应的反粒子，这些反粒子拥有相同的质量和相反的电荷。电子的反粒子叫做正子（positron），而在狄拉克预测其存在之前，人们对它一无所知。

如果把所有的粒子都替换成其反粒子，那么物理定律（几乎）是不变的，这样，就产生了一个和自然界对称的世界。狄拉克从来没有太留意群论，却发现了自然界中最迷人的对称群。

自 1935 年开始，直到 1989 年在塔拉哈西（Tallahassee）去世，狄拉克一直高度重视数学美在物理理论中的作用，并把这一原则作为检验自己研究的试金石。他认为如果一项研究结果不美，就是错的。

1956年他去了莫斯科大学，遵循这所学校的一贯传统，他要为后人在黑板上写下一句格言，他写道：“物理定律必须具有数学美。”他还谈到过自然界的“数学性”（mathematical quality）。但是，他仿佛从来都不认为群论是美的，这可能是因为大多数物理学家都是通过庞杂的推论来了解群论的。好像只有数学家才能理解李群那种精妙的美。

*

不管美还是不美，群论还是很快就成了量子物理学家们的必修课，这件事要归功于一个皮革商人的儿子。

19世纪末，皮革是个大生意，现在也是这种状况。在当时，一个小商人通过鞣制和贩卖皮革就可以过上不错的生活。制革厂厂主安塔尔·维格纳就是一个很好的例子。他和妻子艾尔兹塞贝特都是犹太人后裔，实际上却不是犹太主义者。他们生活在当时属于的奥匈帝国的佩斯城，后来，它与毗邻的布达合并，成了现在的布达佩斯，匈牙利的首都。

他们的第二个儿子耶诺·帕尔·维格纳（Jenő Pál Wigner）生于1902年，5岁到10岁之间一直在家里跟着家庭教师学习。耶诺开始上学不久，就被诊断患上了结核病，并被送到奥地利一家疗养院进行恢复治疗。他在那儿待了六个月，却发现原来自己被误诊了。然而如果诊断是正确的，几乎可以肯定他活不到成年。

由于整天躺着，这个男孩就通过做数学题来打发时间。“我整天都被逼着躺在露台上的椅子上，”他写道，“努力作出给出了三条高线的三角形。”高线就是从三角形的一个角出发，与对边垂直相交的线。给出三个角，作出三条高线，是比较容易的，然而反过来作就困难得多了。

离开疗养院后，耶诺还在思考数学问题。1915 年，在路德高中（Lutheran High School），他结识了一个男孩，这个男孩就是未来世界上最了不起的数学家之一，约翰·冯·诺伊曼（Janós [后改名 Jhon] von Neumann）。但两个人好像一直都是泛泛之交，因为冯·诺伊曼喜欢一个人独处。

1919 年，社会主义者夺取匈牙利政权，维格纳一家逃亡到了奥地利。后来，就在同一年，社会主义者被赶出匈牙利，维格纳一家又回到了布达佩斯。全家人都改信了路德教，但这件事对耶诺几乎没有任何影响，他后来说，这是因为自己“只是个温和的信教者”。

1920 年，耶诺以差不多全班第一的成绩完成了学业。他想当一名物理学家，但是他的父亲想要他加入家庭的皮革经营事业。所以耶诺没能去学物理学，而是学了化学工程学，他的父亲认为这也许对自己的事业有所助益。大学的第一年，维格纳去了布达佩斯技术学院（Budapest Technical Institute），后来就转学去了柏林的高等技术学校（Technische Hochsechule）。他把大部分时间花在了化学实验室里，只有很少的时间花在了理论研究上。

但是耶诺并没有放弃物理学的学习。他离柏林大学不远，那里正是普朗克、爱因斯坦和另外一些不太著名的人物的天下。耶诺利用这种便利，去听了几次不朽的讲座。他以一篇研究分子构成和解体的论文完成了自己的博士学业，并正式进入了制革厂。但这不是什么好兆头：“我在制革厂并不顺心……我在那儿并不得心应手……我觉得这不是我的人生。”他的人生应该是数学和物理学的人生。

1926 年，他和恺撒·威廉研究中心（Kaiser Wilhelm Institute）一位检晶科学家签了合同，做了他的研究助手。这项工作虽然也是化学研究，但却把维格纳的两种兴趣都囊括在内。这项研究对维格纳的

人生和核子物理都产生了巨大的影响，并把他引向了群论——数学对称。物理学对群论的第一次重要运用，是对 230 种可能存在的晶体结构进行分类。维格纳写道："我收到一位检晶科学家的来信，他想知道原子在晶格中的分布为什么是轴对称的。他还告诉我这个问题只能用群论来解决，我必须读一本关于群论的书，然后把研究结果告诉他。"

可能正是因为对儿子在皮革生意中的表现让自己大失所望，安塔尔才同意了儿子这份助手的差事。开始时，耶诺读了一些海森堡关于量子理论的著作，掌握了一种推算带有三个电子的原子光谱的理论方法。但是，他发现在超过三个电子时，这种方法非常复杂。带着这个问题，他向老相识冯·诺依曼求教，冯·诺依曼建议他学学群表示论。群表示论中有很多代数概念和技巧，特别是矩阵代数方面的内容。但是，维格纳研究过晶体学，对一些当时的前沿数学著作（海因里希·韦伯 [Heinrich Weber] 的《代数教程》[*Lehrbuch der Algebra*]）也比较熟悉，所以矩阵对他来说并不是什么障碍。

冯·诺依曼的建议取得了立竿见影的效果。如果一个原子拥有一定数目的电子，由于所有的电子都是相同的，所以就很难"知道"这些电子到底哪个是哪个。换句话说，无论这些电子怎么排列组合，描述原子辐射的方程都必须是对称的。通过群论，维格纳提出了一种拥有一定数目的电子的原子的光谱理论。

从这一点上说，他的研究已经涉及了经典物理学领域。但是，维格纳感兴趣的是量子物理学，他已经开始着手自己毕生的伟大研究——群论在量子力学中的应用。

吊诡的是，他并不是因为自己的新工作才走上了这条研究道路。德国数学界的元老大卫·希尔伯特对量子物理中的数学原理很感兴趣，并且想找一个这方面的研究助手。1927 年，维格纳去了哥廷根大学，

加入了希尔伯特的研究团队。他的职责是为希尔伯特的数学见解提供物理学解释。

但是这项研究进行得并不顺利。希尔伯特年迈、陈腐而孤僻，两个人在一年中总共见了五次面。所以维格纳又回到了柏林，做一些量子物理学方面的讲座，继续创作自己最著名的著作——《群论及其在原子光谱量子力学中的应用》（*Group Theory and Its Application to Quantum Mechanics of Atomic Spectra*）。

在有些方面，他曾被赫尔曼·外尔超越。外尔也写过一本关于量子理论中的群的著作，但是他关注的主要是基础问题，而维格纳则是专门针对物理学问题的。外尔寻求的是美，维格纳追寻的是真。

*

我们可以通过鼓的震动这个简单的例子来解释维格纳的群论思想。鼓一般是圆的，但是原则上讲，它可以是任何形状的。当你用鼓槌敲鼓，鼓面震动，就发出了声音。不同形状的鼓发出的声音也是不同的。鼓发声的频率范围叫音域频谱（spectrum），它对鼓的形状有一种复杂的依赖关系。如果鼓是对称的，其音域频谱也就应该是对称的。没错，其音域频谱是对称的，但是这种对称并不容易把握。

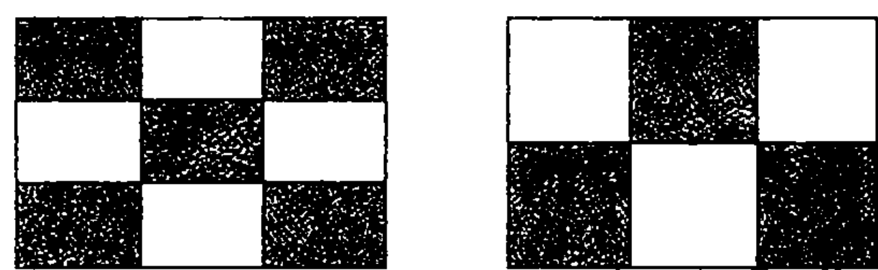
可以想像一个长方形的鼓，除了在数学系里，很难看到这种鼓。这种鼓的典型振动模式，是把鼓面分成若干个小长方形，比如：



长方形鼓的两种振动模式

在这里，我们可以看到两种不同的震动图形，其振动频率也是不同的。上图是这些振动图形的粗略印象，阴影区是向下运动的，空白区是向上运动的。

鼓的对称会对其振动图形产生影响，因为鼓的所有对称转换都会影响不同的振动图形之间的转换。所以振动模式都有对称性。但是各个振动图形的对称并不一定和鼓的对称相同。比如，长方形是180°旋转对称的，如果你把这种对称转换应用到上面两种振动图形中，它们就变成了：



把相同的模式旋转180°

左边的图形没有变化，所以它和鼓具有相同的旋转对称。但是右边图形中的阴影部分和空白部分颠倒了。这种效应叫做"自发性对称破缺"（spontaneous symmetry-breaking），在物理系统中很常见。左边的图形的对称没有受到破坏，但是右边的图形则破坏了。让我们看看是什么破坏了右边的图形的对称。

尽管这个图形和旋转后是不同的，但是它们的频率是相同的，因为旋转既是鼓的对称，也是描述其震动的方程的对称。因此这一频率会在鼓的声音频谱中出现"两次"。看起来，想用实验侦测到这种效应并不容易，但是如果将鼓做一些细小的变动（比如在鼓的边缘刻一条痕迹）两个频率就会发生细小的偏差，你会发现两个非常接近的频率。如果这个频率在对称的鼓的声音频谱中只出现一次，就不会发生这样的现象。

维格纳发现分子、原子和原子核中也会出现这种效应。这里，鼓的声音换成了分子的震动，声音频谱换成了光谱。在量子世界中，光谱是通过不同能量态之间的变换产生的，原子会放射出光子，这些光子的能量会随着能量态的变化而变化。这样，光谱就可以用分光镜进行观测了。同样，由于分子、原子或原子核的对称，有一些频率（我们观测到的是相应的光谱线）是双的，也可能是好几个。

怎样才能确定一个频率出现的次数呢？我们不能像在鼓上刻标记那样在分子上也刻上标记，但是我们可以把分子放在一个磁场中。这样也可以破坏其潜在对称，使光谱线分离。这样，你就可以运用群论（严格地说，是群表示论）来推算频率并研究它是怎样分离的了。

表示论是最美、最强大的数学理论之一，但它也有很高的技巧要求，而且布满了陷阱。维格纳把它变成了一门高超的艺术，其他人则只有被他带着跑的份儿。

*

1930 年，维格纳在美国高等研究院（Institute for Advanced Study）的兼职工作固定了下来，在普林斯顿和柏林之间来回奔走。1933 年，纳粹通过了禁止犹太人在大学任教的法令，所以维格纳永久性地移居到了美国，主要待在普林斯顿，他还把自己的名字英语化为尤根·保罗（Eugene Paul）。他的妹妹马尔基特也到了普林斯顿，和他生活在一起。在那儿，她遇见了正在访学的狄拉克，1937 年，两个人结了婚，所有人都为此大吃一惊。

马尔基特的婚姻很美满，而尤根的工作却很不如意。1936 年，维格纳写道：“普林斯顿大学把我解雇了，他们从来没有解释过为什么

要这么做。我不禁为此感到气愤。”事实上，是维格纳辞职的，而且很明显是因为他觉得提升不够快。也许他认为普林斯顿大学不愿意提升他是为了迫使他辞职，所以才觉得自己被炒了鱿鱼。

他很快在威斯康星大学找到了新工作，并获得了美国籍，还碰上了物理学学生艾米利亚·弗兰克（Amelia Frank）。他们结婚了，但是罹患癌症的艾米利亚不到一年就去世了。

在威斯康星，维格纳把注意力转移到了原子能上，并发现它受控于对称群 SU(4)。他还作出了一个关于洛伦茨群的基础性发现，发表于1939年。但是在当时，群论还没有成为物理学家正规训练中的一部分，它的主要应用领域仍然集中在晶体学中。对大多数物理学家而言，群论看起来既复杂又陌生，简直是要命。量子物理学家们被闯入自己园地的群论吓坏了，他们称其为“群祸”（Gruppenpest），或者“群病”（group disease）。维格纳已经开启了这一风气，但他的同事们却不愿意跟上。但是维格纳的见解极具前瞻性。因为对称的影响无孔不入，群论法最终主导了量子力学。

1941年，维格纳开始了自己的第二段婚姻，妻子是一个名叫玛丽·安妮特（Mary Annette）的老师。他们生了两个孩子，大卫和玛莎。第二次世界大战期间，维格纳和冯·诺依曼等一大批顶尖级数学物理学家都参与了制造原子弹的曼哈顿计划。他于1963年获得了诺贝尔物理学奖。

尽管在美国生活了很多年，维格纳还是一直思念着家乡。“在美国生活了60年，”他在耄耋之年写道，“我仍然更像个匈牙利人，而不是美国人。很多美国文化和我格格不入。”他死于1995年。物理学家亚伯拉罕·派斯这样形容维格纳：“一个不寻常的人……20世纪的物理学巨人。”他打开的视野在21世纪仍然是革命性的。

五维的人

20 世纪后期，物理学取得了惊人的进展。广义相对论已经很好地描述了宇宙的大尺度结构。黑洞（由大质量的恒星在自己重力的作用下坍塌而形成的时空范围，光都无法从那里逃脱）之类著名的预测，已经得到了观测资料的支撑。另一方面，现代形式的量子理论（量子场理论）和狭义相对论一起，详尽而精确地描述了宇宙的小尺度结构。

但是，科学家的天堂里还有两条蛇*。一条是“哲学性的”蛇：这两种取得巨大成功的理论是互相抵触的，他们对物理世界的设想是不一致的。广义相对论是“决定论的”——它的方程不容忍任何随机性；量子理论却有一种固有的非决定性，其中包括海森堡的测不准原理，还有一些现象，比如放射性原子的衰变，也是随机发生的。另一

---

* 蛇在基督教教义中是反上帝的，正与天堂相对。在伊甸园中，引诱夏娃犯禁的正是蛇。在这里指的应该是危险。

条蛇是“物理的”：关于基本粒子的量子理论中还有好几个悬而未决的问题，比如，粒子为什么具有一定的质量，或者干脆就问它为什么具有质量。

很多物理学家都相信，有一个大胆的办法可以把这两条蛇通通赶出伊甸园：统一相对论与量子理论。也就是说，在逻辑统一的前提下发明一个新理论，并使其在大尺度上符合相对论，同时在小尺度上符合量子理论。这正是爱因斯坦在后半生所做的事，但是他失败了。科学家们抱着一贯的谦恭，将这种理论命名为“万有理论”。他们希望能把整个物理学归纳为一个可以写在T恤衫上的方程。

这并不是异想天开。你完全可以把麦克斯韦方程写在T恤衫上，我就有一件印有狭义相对论方程的T恤衫，上面还用希伯来语写着旧约中的名言“要有光”，是我的一个朋友在特拉维夫机场买给我的。说回正题，此前已经有很多把迥然不同的理论统一在一起的成功范例了。麦克斯韦的理论就统一了电和磁，而人们曾将它们看成两种完全不同的自然力所产生的毫不相干的现象，结果却变成了一个独立的现象：电磁。名字也许比较蹩脚，但却传神地反映了统一的过程。还有一个更晚近的例子，但除了物理学界之外，没多少人知道，那就是电弱理论（electroweak）。它是由电磁与弱核力（weak nuclear force）统一而来的，我们下面还会讲到。电磁和强核力的统一一直未能实现，它们的统一结果将是：重力。

虽然这一统一没能实现，但我们完全有理由设想，这种终极自然力会将物理学推向终结。不幸的是，重力的一些棘手特性让这一进程变得十分困难。

*

也许根本就没有什么“万有理论”。尽管作为“自然法则”的数学方程已经这样成功地解释了我们所在的世界，但仍然不能保证它一直都会这么成功下去。也许，宇宙根本没有物理学家想像得那么具有数学性。

数学理论可以与自然非常近似，但并不一定在每一点上都与之完全一致。如果是这样，就需要对相互抵触的理论进行修补，以获得可以同时适用于不同领域的近似性。并且，也不可能存在一种把所有这些近似性杂糅起来的至高无上的理论。

当然，还是有这样一种毫无意义的法则：“如果速度慢，尺度大，就采用牛顿力学；如果速度快，尺度大，就采用狭义相对论，”如此等等。这种混搭理论丑陋得可怕。如果美就是真，那么这种混搭只能是错误的。但是，也许宇宙从根源上来说就是丑的；也可能宇宙根本就没有什么所谓根源。话说起来也许不好听，但是，我们有什么资格把自己狭隘的美学强加在宇宙之上呢？

认为“万有理论”必然存在的观点让人想起一神教。在一神教中，一千年来各管一摊的神祇被一个掌管一切的神所取代。人们普遍认为这一过程是一种进步，但这和一种“未知数的方程”式的哲学错误很相似，把同一个原因摊派给所有的神秘现象。正如科幻小说家伊萨克·阿西莫夫（Issac Asimov）所言，如果你对飞碟、心灵感应和幽灵都不了解，那么给你的最好“解释”就是：一个会心灵感应的幽灵驾驶着飞碟。这种“解释”给人一种这样的错误认识：我们从前有三个需要解释的问题，现在只剩下一个了。但是剩下这一个问题是原来三个各自有着不同解释的问题统一而成的。通过这种统一，我们将其

可能性抛在了脑后。

当你用太阳神解释太阳、用雨神解释雨的时候，你可以赋予每个神不同的特征。但是如果你坚持认为太阳和雨都是由同一个神掌管的，你就得把两种不同的东西塞到同一个套子里。因此，从某种意义上说，基础物理学就是原教旨主义物理学。只不过，内在的神换成了可以印在T恤衫上的方程，方程的展开代替了涉入世间的圣徒。

尽管有所保留，我还是赞成物理学的原教旨主义。我希望能够看到一种万有理论，如果在数学上，它既是美的，又是真的，我会更加高兴。我想信教的人们也会很高兴，因为这证明他们所信仰的神具有高妙的趣味和智性。

*

当今对于万有理论的蕲求发端于统一电磁学和广义相对论的尝试，在当时，这二者就是物理学的全部。这种尝试是在爱因斯坦狭义相对论论文发表了14年后，也就是光会因重力而弯曲的预测作出的八年后，广义相对论提出的四年后开始的。这是一种非常有益的尝试，它会将物理学变成一个整一的新课题；但是很可惜，爱因斯坦的研究反倒和另一个新课题——量子力学挂上了钩。物理学家们争先恐后地到量子力学领域去淘金，逐渐对统一场论失去了兴趣，因为量子物理学中更容易作出重大发现。从最初的尝试到这一思想的诞生，一共花了60年的时间。

这一思想肇始于柯尼斯堡，当时是德国东普鲁士省的首府。柯尼斯堡就是现在的加里宁格勒（Kaliningrad），是位于波兰和立陶宛之间的俄罗斯行政中心。这座城市对数学发展史的影响是从一个难题

开始的。这座城市坐落于普雷格尔河（Pregel）沿岸，有七座桥连结着两岸和中间的两座小岛。有没有一条让柯尼斯堡的市民依次经过每座桥而又不重复经过任何一座桥的路线呢？有一位叫莱昂纳德·欧拉（Leonhard Euler）的市民提出了一种关于这种问题的普遍理论，他认为不可能有这样一条路线。这种理论向数学拓扑学领域迈出了第一步。拓扑学研究的是一个图形受到折弯、扭曲、挤压等各种不影响其连续性的变形（不包括切割和拆散）时，仍然保持不变的几何性质。

拓扑学已经变成了最强大的数学工具之一，在物理学中应用非常广泛。它能告诉我们多维空间中可能存在的形状，这在宇宙学和粒子物理学中都变成了越来越重要的课题。在宇宙学中，我们可以通过拓扑学来研究最大尺度（整个宇宙的尺度）上的时空形状；在粒子物理学中，我们可以通过它来研究小尺度上的时空形状。你可能会觉得答案显而易见，但是物理学家却不这样认为，这种怀疑也是从柯尼斯堡开始的。

1919年，柯尼斯堡大学籍籍无名的数学家提奥多·卡鲁查（Theodor Kaluza）产生了一种古怪的想法。他把自己的想法写信告诉了爱因斯坦，而爱因斯坦差点吃惊得说不出话来。卡鲁查找到了一种融合重力和电磁的方法，那就是“统一场论”，而爱因斯坦已经在这一问题上摸索多年，但是一无所获。卡鲁查的理论非常简洁，也非常合理，但是只有一个麻烦：要实现二者的统一，时空必须是五维的。时间的维度不变，但空间却要多出一个第四维。

卡鲁查并没有着手去统一重力和电磁。由于某些他认为非常重要的原因，卡鲁查围着五维重力问题兜起了圈子。五维重力问题只是数学家的热身练习，目的是弄清在存在未知维度的情况下，爱因斯坦场方程会出现怎样的情况。

在四维空间中，爱因斯坦的方程有 10 个“成分”（component），可以分解成 10 个描述 10 个不同的数的独立方程。把这些数放在一起，可以描述表示空间曲度的张量。在五维情况下，爱因斯坦的场方程有 15 个成分，因此也就有 15 个方程。其中 10 个构成了标准的爱因斯坦四维理论，并无什么特别之处；四维时空内嵌于五维时空之中，因此你可能会很自然地猜到四维重力也内嵌于五维重力之中。而另外的五个方程呢？它们可能描述了与我们的世界无关的某种特殊结构。但事实并非如此。我们对它们非常熟悉，而正是这一点，让爱因斯坦大为吃惊。卡鲁查剩下的方程中，有四个正好是麦克斯韦的电磁场方程，并没有超出我们的四维时空。

剩下的最后一个方程描述了一种非常简单而且无足轻重的粒子。但是任何人（尤其是卡鲁查）都没有料到，爱因斯坦的重力理论和麦克斯韦的电磁学理论在五维重力的情况下自动呈现。卡鲁查的推论似乎意味着，光是一种在另外一个隐秘的空间维度中的震动。你可以把重力和电磁放在一个浑整的统一体中，前提是，空间必须是四维的，而时空则必须是五维的。

卡鲁查的论文让爱因斯坦非常头疼，因为没有理由让时空凭空多出一个维度来。尽管这种思想十分古怪，但是由于它很优美，而且很可能对将来的研究产生重大影响，爱因斯坦还是决定将它发表。在斟酌了两年之后，爱因斯坦把这篇论文发表在了一个重要的物理学刊物上。题目是“物理学问题的统一”。

*

以下对于新增的时空维度的讨论可能过于模糊和神秘了。但是，

这和维多利亚时代的气质是分不开的，人们会把第四维度当做解释所有未知事物的万灵药。灵魂在哪儿？在第四维。灵性从哪儿来？第四维。信众们甚至把上帝和天使都放在了第四维中，后来他们又发现第五维更好，再后来，他们又相中了第六维，就像狗熊掰棒子一样。到头来，这些全知全在的实体只能跑到一个无限高的维度中去了。

这些说法都很有趣，但是太不科学了。所以，必须弄清楚它所包含的数学内容。最重要的一点是，我们必须清楚，“维度”指的都是描述数学和物理学论题的不同变量。

科学家对变量作出过大量的思考，实验物理学家甚至对其作出过很多次的测量。“维度”只是变量的几何式说法，事实证明，这种说法非常有用，以至于它已经深深嵌入了数学和物理学中，人们对此已经习以为常，因此也不太在意它是怎么回事了。

时间是一个非空间变量，因此也就变成了一个“可能的”第四维度。但是，温度、风俗和坦桑尼亚白蚁的生命周期也都可以成为一个维度。一个三维空间中的点是由三个变量决定的，其中包括相对于参照点的东西方向上的位置、南北方向上的位置和上下方向的位置，并用负数表示相反的方向。以此类推，四维空间中的所有事物都是由四个变量决定的，而 101 维中的所有事物都是由 101 个变量决定的。

有很多复杂的系统本身就是多维的。比如说你院子里的天气情况就取决于温度、湿度、风速的三个要素、大气压和降雨强度，总共有七个维度，而且我们还排除了很多其他的维度。我打赌我从没意识到自己有一个七维的院子。太阳系九个行星（就算是八个吧——哎，可怜的冥王星）的状态是由每个行星的六个变量决定的，其中包括三个位置参数和三个速度变量。所以我们的太阳系在数学上是 54 维的，这还没有算上卫星和小行星。一个经济系统都有上百万种不同的商

品，每种商品都有自己的价格，所以经济系统就有上百万个维度。在电磁学中，确定一定区域的电磁场的状态只需六个数字，真是小巫见大巫。类似这样的例子还有很多。当人们开始科学地研究具有巨多变量的系统时，就必然牵涉到具有超多维度的空间。

关于多维空间的正式数学是纯代数性的，而且很明显只是把低维空间（low-dimensional）推而广之罢了。平面（二维空间）上的点可以用两个参数来确定，每个三维空间中的点都可以用三个参数来确定。用四个参数来确定四维空间中的点也没什么稀奇，确切地说，用 n 个参数确定 n 维空间中的点也没什么了不起。n 维空间本身就是所有这些点的集合。

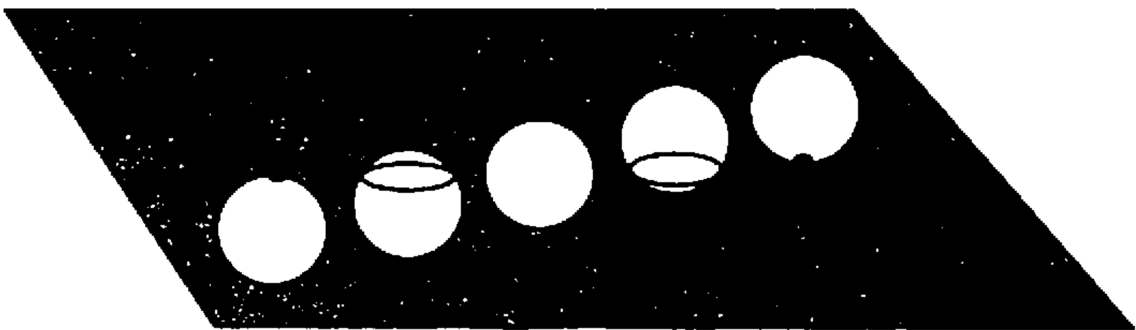
通过类似的代数运算，你可以算出 n 维空间中的两个点的距离、两条线的角度，不一而足。然而接下来要做的，就得借助想像力了：大多数二维空间和三维空间中的几何形体都和 n 维空间中的情况有着明显的相似性，找出这些相似性的方法，就是先用代数参数描述二维空间和三维空间中的形体，再描述 n 维空间中的相似形体。

*

想要获得对 n 维空间的感性认识，必须借助关于 n 维空间的可视图景。而要获得其视觉形象，我们就得跟英国牧师兼校长埃德温·A.艾伯特（Edwin Abbott Abbott）学两招。1884 年他写过一本名为《平面世界》（*Flatland*）的小书。这本书写的是 A. 平方（A.Square）的冒险故事，这个 A. 平方先生就生活在二维的欧几里得平面上。艾伯特并没有告诉我们“A”代表的是什么，我认为应该是“Albert”，个中原因我已经在我为这本书写的续编《更平的世界》（*Flatterland*）中阐

述过了。脑袋少根弦的阿尔伯特·平方不相信有什么三维空间，直到有一天，发生了一件天大的事情，一个球体穿透了他的平面世界，并把他抛进了自己从未想像到的三维空间中。

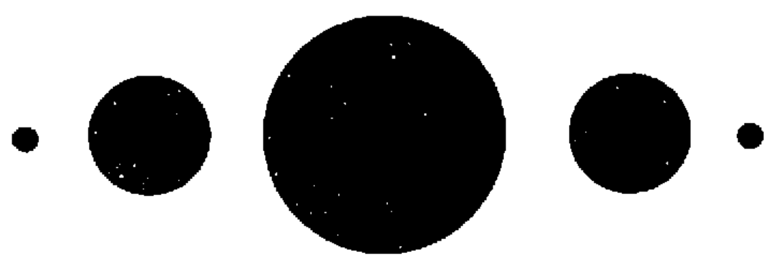
《平面世界》借寓言的形式，讽刺了维多利亚社会通过跨维度类推产生的关于第四维度的观念。而我们这儿要讲的不是讽刺，而是类推。如果能够想像你自己在二维空间中，傻呵呵地不相信三维空间的存在；那么想像一下你在三维空间中，却还是傻呵呵地不知道四维空间的存在，就也不是什么难事。我们可以这样设想，阿尔伯特·平方先生整日坐在平面上，想要“想像出”一个实在的球体。通过让一个球体穿透平面世界，艾伯特成功地做到了这一点。由于球体是垂直穿过平面的，所以阿尔伯特也就看到了它的横断面。最初，他只看见一个点，慢慢变成一个圆盘。他看见圆盘时，正是这个球体的赤道（最大圆周）穿过平面的时候，接着，圆盘开始缩小，最后消失无踪。



球体穿过平面世界

事实上，阿尔伯特看到的这个圆盘是一些逐次进入阴影的线，但他把自己的看到的东西想像成了一个盘子，这和我们看见一个面时也会把它想像成某种立体的东西是一个道理。

以此类推，我们可以“看到”一个“超球面”（hypersphere），它是从三维空间的球体类推出来的。它也是从一个点开始，慢慢扩大，直至我们看到其“赤道”，接着就逐渐缩小成一个点，然后消失。



超球体穿过三维世界

空间真的不止三个维度吗？我们这里说的并不是非时间维度的数学构想，而是真正的物理空间。归根结底，第四维度是怎么填满的呢？该填满的早已经填满了。

如果你这样想，就说明你没有从阿尔伯特·平方那儿吸取教训，他对平面的看法与此如出一辙。抛开浅薄的偏见，我们就可以想像，原则上讲，空间完全可以是四维的，或者是上百万维的，如此不一而足。但是，日常的观察告诉我们，在我们这个特定的世界中，上帝给了空间三个维度，还为时间留了一个维度。

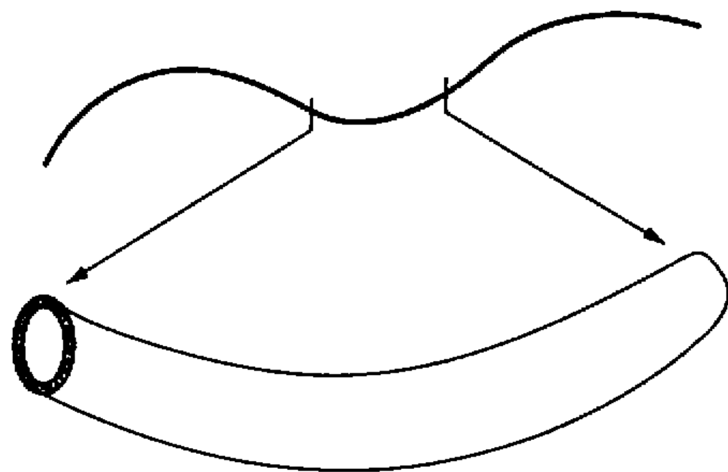
上帝是这样做的吗？科学家给我们的教谕只有一条，那就是要对日常观察提高警惕。一把椅子仿佛是密实的，但它绝大部分是空的。空间看似平直，但却有着一定的曲度。量子物理学家认为，在非常小的尺度上，空间是一些量子泡沫（quantum foam），绝大部分是孔穴。根据量子物理的测不准原理提出的“多重世界”理论的支持者们则认为，我们的宇宙只是无限多种共在宇宙中的一个，而我们只是多重宇宙中的沧海一粟。如果常识会在其他方面误导我们，那它在空间和时间的维度问题上也会误导我们。

*

卡鲁查对自己时空理论中多出的维度作出了一个简单的解释。传

统的维度都是直线方向的，而且可以延伸到很远，其长度可达到几十亿光年。卡鲁查这个新维度却绝非如此：它卷成了一个圆周，而且比原子还小。构成光波的脉动可以绕着它运动，因为它也比原子小；但是物质不能在其中运动，因为它内部的空间太小了。

这种思想绝不是荒唐可笑的。如果你站在远处看一条橡胶软管，这个软管看起来就只是一条曲线，而曲线只有一个维度。只有离得很近的时候，我们才会发现这个软管原来是三维的，有一个二维的横断面。这个我们在远处看时没有注意到的横断面结构，正是使软管可以输送水流的原因。这个横断面需要有相应的形状，以及一个位于中间的孔。我们可以想像一下，这条橡胶管的粗细还容不下一个原子。你必须离得异乎寻常地近，才能发现这个新维度。这个细得让人难以置信的橡胶管肯定输不了水了，但是足够小的东西却能够从中通过。



在远处看，一条橡胶管是一维的（上）；靠近看的时候却多出了两个维度（下）。

所以，我们可以通过理解这个新维度产生的效应来理解这个维度。这也就是说，这个隐而不显的维度完全是一个科学设想：它的存在在理论上讲是可以验证的，但是必须通过干扰来进行，而不是直接进行检验。许多科学检验都是通过干扰来进行的——如果你能够直接看到某些现象的原因，那也就用不着理论和试验了。我们可以举个例子，没有人看到过电磁场。人们看到过火花和指南针指向北，于是人

们（如果他们是科学家的话）就可以推论说肯定存在着相应的电磁场。

卡鲁查的理论得到了一定程度的传播，因为它是惟一可以让人们对统一场论抱有希望的理论。1926年，另外一个数学家，奥斯卡·克雷恩（Oskar Klein），又将卡鲁查的理论向前推进了一步，他认为，量子学也许可以解释为什么这个第五维度会卷成如此严密的圆形。事实上，他的大小可能和普朗克常数属于同一量级："普朗克长度"是10^{-35}米。

物理学家们很快被卡鲁查—克雷恩理论吸引了。但是，这个新维度的存在无法直接证明，这让科学家们伤透了脑筋。很明显，卡鲁查-克雷恩理论符合了众所周知的重力和电磁现象。你永远也无法用标准的实验证明它不存在。但是它不会产生什么新的东西，也没有预言什么可以验证的东西。很多试图统一现行规律的努力也因相同的问题让人们想破了头。可以验证的东西都是已知的，新东西都是无法验证的。于是，最初的热情开始减退了。

对于卡鲁查-克雷恩理论的致命打击既不是它的正确性问题，也不是不值得花那么多的时间对其进行研究，而是比它更有吸引力的理论的蓬勃发展，这个理论确实可以作出新的预言，而且可以用实验进行验证。这就是量子理论，它当时正进入第一个迅猛发展的阶段。

*

但是，到了1960年代，量子力学也开始止步不前了，早先的进步被一些深奥的难题和让人摸不着头脑的观测结果所取代。量子理论取得的成功当然不容置疑，并且很快就将建立基本粒子的"标准模型"(standard model)。但是要找到解决问题的新方程，却一度变得非

常困难。真正新颖的思想很难检验，而可以检验的思想则往往只是某种现存思想的延伸。

所有的实验都揭示了一个根本原理：解开物质的小尺度结构秘密的钥匙就是对称。基本粒子的对称既不是欧几里得空间中的简单运动，也不是相对论时空中的洛伦茨变换。它包括“规范对称”（gauge symmetry）和“超对称”（supersymmetry）。还有另外一些对称，和加洛瓦研究过的对称相似，也是通过不连贯集合的排列来实现的。

这些对称是怎么回事呢？

对称都是从群中得出的，但是，在很多的“操作”方法中，群并不会产生对称。在诸如旋转、重新组合以及时间反演等操作方法中，群是可以实现对称的。粒子物理学发现了一种实现对称的新方法，叫做“规范对称”。这个名字只是出于历史的偶然，而更好的叫法应该是“局部对称”（local symmetry）。

假设你在另一个国家旅行，我们就叫他“二倍国”*（Duplicatia）好了。二倍国的通行货币是凡尼（pfunnig）** 而汇率是两美元可兑一凡尼。在你正为此烦恼时，却发现了用美元兑换凡尼的法则。这个法则就是，买每样东西所花的美元正好是所花凡尼的两倍。

这就是一种对称。你可以把所有的数字都扩大一倍，但商业交换的“定律”是不变的。为了弥补数字上的变化，你就得用凡尼付款，而不是用美元。这种“货币量改变时的不变性”，就是商业交换规则在全球范围内的对称。如果你从头到尾都进行同样的改变，规则仍然是不变的。

然而，在国界的另一边，还有一个叫“三倍国”（Triplicatia，

---

* duplicate 有二倍、双重的意思，读下文会发现作者的用意。

** 德国有一种货币单位叫 pfennig，即芬尼。

triplicate 有变成三倍之意）的邻国，当地的通行货币是“布德尔”（boodle，赌金，伪币），一布德尔价值 3 美元。如果你要到三倍国去旅行，根据相应的对称，你所带的金额就要乘以 3。但是，相关的商业交换规则仍然是不变的。

现在，我们获得了一种从一个地区到另一个地区的“对称”。在二倍国，这个对称就是乘以 2；而在三倍国，是乘以 3。你很自然地就会想到，到了五倍国（Quintuplicatia），这个乘数就会变成 5。所有这些对称都可以进行类似的运用，但是每一种对称都必须在相应的国家才有效。这些商业法则是不变的，但是对于不同的货币，必须对症下药。

这种货币量兑换的区域性改变就是商业法则中的一种规范对称。原则上讲，汇率会随着时间和地点的变化而变化，但是，法则会保持不变，以便根据“货币场”（currency field）中的区域性货币价值来解释每一桩交易。

*

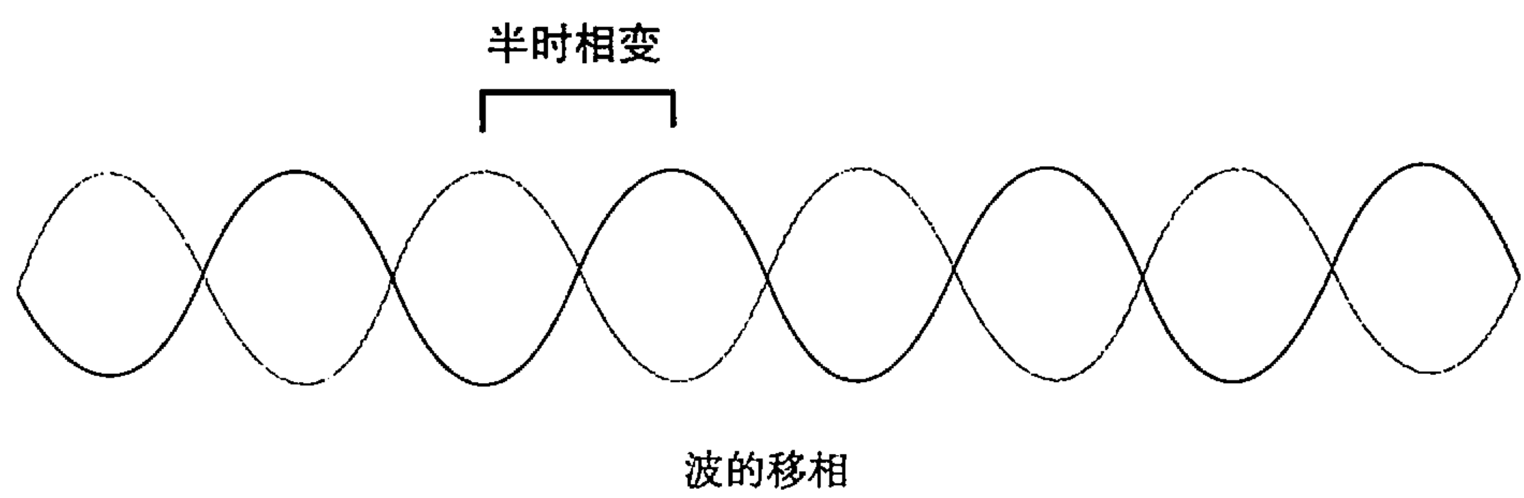
量子电动力学融合了狭义相对论和电磁学。这是麦克斯韦以后的第一次物理学统一，这个统一的基础就是电磁场的规范对称。

我们已经了解到，电磁在狭义相对论的洛伦茨群中是对称的。这个群是由地球空间对称构成的，这也就意味着，如果我们要保留麦克斯韦方程，那它必须是整体对称的。但是，麦克斯韦电磁学也有一个规范对称，它对于量子电动力学至关重要。这个对称就是光的“相变”。

所有的波都是由有规律的摆动构成的，最大的摆动幅度就是波的

振幅。完成一个最大摆动幅度的时间点叫做“相位”（phase），它告诉我们波峰会出现于何时、何处。了解某种特定的波的相位没有多大的意义，重要的是两种不同的波的相位差。比如说，如果两个相同的波是半周期（两个波峰出现的时间的中间点）关系，那么其中一个波的波峰出现的时间点就和另一个波不一致，而是与其波谷出现的时间点一致。

在街上走路时，你的左脚和右脚就是半周期关系。如果是一只大象在街上走路，那么它的四只脚触地的相位就分别位于全周期的0，1/4，1/2和3/4处，首先是左后脚，再是左前脚，接着是右后脚，最后是右前脚。你可以这样来理解，如果们从0开始数，每只脚会对应一个不同的数，但是这些相位差仍然是0，1/4，1/2和3/4。这样，相对相位就得到了明确的界定，在物理学上也就有得讲了。



波的移相

我们可以假设一束光穿过一套由透镜和反光镜组成的复杂系统。它的传播和总体相位的关系并不明显。相位的改变就相当于将观察往后延一小段时间或观察者重置自己的钟表，它既不会影响对称系统的几何性质，也不影响光的传播路径。甚至有时即使两道光是重叠的，但只要两个波的移相数量相同，照样不会有任何变化。

到现在为止，“变相”仍然是整体对称。但是，如果在仙女星系（Andromeda galaxy）的某处，有一位外星实验家在自己的试验中改变

了光的相位，我们在地球上任何一个实验室中都是不会感受到任何影响的。所以，在任何地方，光的相位都是可以随意改变的，而物理定律却不会改变。在时空中的任一点，光都具有发生相变的可能性，并不一定要求非得是全局性的变化不可。这种可能性是麦克斯韦方程的一种规范对称，量子电动力学方程中也有这种对称。

一个全周期移相（phase shift）和不进行移相完全没有区别，在抽象意义上讲，这也预示着相位改变就是一种旋转。所以，这里牵涉到的对称群是二维旋转群 SO（2)。但是，物理学家们希望自己的量子参数的变换是“酉性的”（unitary，由复数限定，而不是实数)。幸好,SO（2）群中有一个典型的酉性群 U（1)，可表示复平面（complex plane）中的旋转。

总而言之：量子电动力学具有 U（1）的规范对称性。

规范对称是下面要讲到的两种物理学统一——电弱理论和量子色动力学（chromodynamics）——的一条线索。这构成了量子物理的“标准模型”，也就是现在受到普遍接受的基础粒子理论。在讲述这一统一之前，我们必须明白：要统一的不是理论，而是力。

*

当今，物理学已经发现了四种不同自然力：重力、电磁力、弱核力和强核力。这四种力各自有着不同的特点：它们发生作用的时空尺度不同，有的使粒子相互吸引；有的使粒子相互排斥；有的根据粒子的不同，有时使其相互吸引，有时又使其相互排斥；有的则是根据粒子之间距离的不同来决定其相互吸引还是相互排斥。

乍看起来，这四种力彼此之间几乎没有任何相似性。但是，这只

是表象而已，事实上，它们之间的相似性比区别性大得多。物理学家们已经将关于它们之间深层的统一性的证据整理出来，他们认为这四种力有一种通用的解释。

我们时时刻刻都受着重力的影响。当我们把一只盘子掉到厨房地板上，盘子摔成了碎片，我们看见重力把它牵向地球中心，而地板却挡住了它的去路。由于电磁力的作用，冰箱门上的塑料猪才得以不掉下来，而这只是电磁力的一个方面，麦克斯韦当年展示的就是电磁力的这一方面。另一面是电力，冰箱的工作就必须借助电力。其实，摔碎的盘子也揭示了电磁力的作用，只是不太明显，因为其中起着主要作用的是使物体结为一体的化学键。当盘子的拉力超过了使分子聚集在一起的电磁力，盘子就破了。

其余的两种自然力是原子核层面的力，不是那么容易发现。但是，如果没有它们，就根本不可能存在什么物质，因为是它们把原子聚集在一起的。这就是盘子、冰箱、地板，以及厨房得以存在的原因。

原则上讲，其他类型的力可能会产生其他类型的宇宙，而我们几乎完全忽视了这种可能性。人们常常认为，如果我们失去了某些特定的力，生命就会不存在，并以此来证明我们的宇宙非常适合生命的发展。这种观点纯属臆想，是对生命构成的狭隘认识的夸张表现。宇宙中也许不会再有我们这种类型的生命，但是要认为我们这种类型的生命是惟一的生命类型，那就太自负了。人们的错误在于混淆了生命的充分条件（我们的宇宙中那些我们赖以存在的因素）和必要条件。

在这四种力中，第一个得到科学阐释的是重力。就像牛顿所观察到的，这是一种引力：宇宙中的任何两个粒子都会因万有引力的作用

而相互吸引。重力是长程的（long-range），它会随着距离的增大而减弱，但是减弱速度非常慢。从另一方面讲，重力比另外的三种力都弱得多。一个很小的磁铁就可以把一个塑料猪牢牢地固定在冰箱上，整个地球的重力都无法把它拉下来。

第二个独立出来的力是电磁力，由于它的作用，粒子会相互吸引或相互排斥。决定它们是吸引还是排斥的，是它们所带的电荷是否相同，或它们的磁极是否相同。如果是相同的，就会产生斥力；如果是不同的，就会产生引力。电磁力也是长程力。

原子核是由比它更小的粒子——质子和中子（neutron）构成的。由中子的名字就可以知道，他不带电荷，但是质子却带有正电荷。质子之间的电磁斥力会使原子核发生爆炸。是什么把它们聚在一起的呢？重力太微弱了，想想前面的塑料猪你就明白了。把他们聚在一起的一定是另一种力，这就是科学家们说的强核力。

但是如果这种力可以战胜电磁斥力，那为什么宇宙中所有的质子没有全都吸到一起，从而变成一个巨大的原子核呢？很明显，这种强核力在距离增大到原子的大小时，就很快消失了。所以强核力是短程力。

但是强核力解释不了放射性衰变（radioactive decay）。在放射性衰变中，某种元素的原子会“吐出”粒子和射线，并变成不同的元素。比如，铀是放射性的，最后会变成铅。所以肯定还有一种亚原子力。这就是弱核力，与强核力相比，它就是更短程的力，只在原子大小的一千分之一的距离内发生作用。

*

在人们得知物质世界的大厦只是由质子、中子和电子建成的之后，物理学就变得简单得多了。这些“基础粒子”（elementary partical）就是原子的部件，尽管“原子”的意思是“不可再分的”，但它还是可以拆分和切割的。在尼尔斯·玻尔的早期模型中，原子看上去是由质子、中子，以及在一定的距离内绕着它们旋转的、比它们更轻的电子组成的。质子带有一定量的正电荷，而电子带有数量相等的负电荷，而中子则是不带电的。

后来，随着量子理论的发展，这个太阳系式的想像被一种更精细的想像所取代。电子不再被是绕着原子核转动的明确的粒子，而是模糊成了一片有趣的云。这些云被解释为概率云，而这个解释应该是最准确的了。如果你要寻找电子，就到云最浓的区域去找，云稀薄的区域则很难找到。

物理学家们又发明了一种研究原子的新方法，把它分成一片一片的，然后研究每一片的内部结构。一个主要的方法就是，用另一个原子或粒子撞击它，然后观察会飞出什么，这种方法现在仍在使用。随着时间的推移（故事的细节非常复杂，我们无法照顾到细节），人们发现了越来越多不同种类的粒子。其中有中微子（neutrino），它可以畅通无阻地穿越数百万英里，因此很难对其进行侦测；还有正电子（positron），它很像电子，但是没有相反的电荷，不符合狄拉克预言的物质/反物质体系。

“基础”粒子的数量猛增到了60种，于是物理学家们开始寻找他们的深层秩序，因为基础性的材料不应该有这么多。每种粒子都可能具有一系列的特点：质量、电荷，以及某种被称为“自旋”（spin）的

东西，因为粒子都是绕着某种轴线（过去，人们一提到旋转就会想到轴线，但这种轴线事实上非属此类）自旋着的。电子并不像地球和陀螺那样在空间中旋转，而是在一个外在维度中“spun”——就让这个单词随便代表什么意思好了。

和量子世界中的所有东西一样，这些性质大多都是一个很小的基数（量子）的整数倍。所有的电荷都是质子所带电荷的整数倍；所有自旋都是电子自旋的整数倍。我们并不清楚质量是否也是这样的，基本粒子的质量问题仍然没有理清头绪。

某种家族相似出现了。必须区分出两种粒子，第一种粒子的自旋是电子自旋的奇数倍，而另一种粒子的自旋是电子自旋的偶数倍。这种区分是根据它们的对称的性质作出的。如果把（在一个外在维度中）自旋放到我们的空间中，就会发生变化。但不管怎么说，这个外在维度好歹算是与我们常见的维度联系起来了。

奇数粒子叫费米子，偶数粒子叫玻色子，分别是以粒子物理学巨擘恩里科·费米（Enrico Fermi）和塞特延德拉·纳特·玻色（Satyendra Nath Bose）的名字命名的。由于某种看似合理的原因，人们一度把电子的自旋值定为 1/2，所以，这样玻色子就有整数个自旋（1/2 的偶数倍是整数），而费米子的自旋数则是 1/2，3/2，5/2，如此等等，相应的，还有 −1/2，−3/2，−5/2 等。费米子遵循泡利不相容原理，这个原理说的是，在任何规定的量子系统中，两种不同的粒子不可能同时处于相同的状态；而玻色子则不遵循这个原理。

费米子包括了所有我们熟悉的粒子，质子、中子和电子全都是费米子。除此之外，费米子还包括牟子（muon）、陶子（tauon）、拉姆达（lambda）、赛（xi）和欧米伽（omega）等更少见的粒子，这些粒子都是用希腊字母命名的。分别与电子、质子和陶子相关的三种中微

子也都是费米子。

玻色子就更神秘了，有着“π 介子”（pion）、“k 介子”（kaon）之类的名字。

现在，粒子物理学家们已经知道了这些粒子的存在，而且可以对其物理特性进行测量。问题是怎么对这堆乱七八糟的粒子作出合理的解释。难道宇宙是抓到什么是什么地随意建成的？或者说，是根据某种尚不为人知的计划建成的？

研究的结果表明，很多原来被认为是基本粒子的粒子其实也是合成的。它们都是由夸克组成的。夸克这个名字是从《芬尼根守灵夜》（*Finnegans Wake*）中来的，夸克有六种“味”（flavor），可以分别叫做：上夸克、下夸克、奇夸克、粲夸克、顶夸克和底夸克。它们都是费米子，自旋数为 1/2，并且各有一个相关的反夸克。

夸克有两种组合形式。其中一个形式是，用三个夸克组成一个费米子。质子由两个上夸克和一个下夸克组成；中子则是两个下夸克和一个上夸克；有一种奇怪的粒子名叫欧米伽负粒子，是由三个奇夸克组成的。另一种组合方式是，用一个夸克和一个反夸克组成玻色子，它们被核力分开，所以不会相互抵消。

要使电荷正常运转，夸克上的电荷数就不能是整数，有些有 1/3 电荷，有的则有 2/3 电荷。夸克还有三种“色”（color），这样就有了 18 种夸克和 18 种反夸克。哦，对了，不止 18 种。因为还要有一些夸克来“携带”弱核力，好把夸克聚集在一起。这样就产生了一种新理论，尽管这个理论使粒子的种类激增，但却具有优美的数学形式。这就是量子色动力理论。

*

量子理论通过粒子交换解释了所有的物理力。就像一只网球把比赛中站在球场两边的两个运动员紧紧联系在一起一样，大量的粒子都带有电磁力、强核力和弱核力。带有电磁力的是质子；带有强核力的是胶子（gluon）；带有弱核力的是中间向量玻色子（intermediate vector boson），它还有个更常见的名字叫"弱子"（weakon，别怪我，这个名字不是我发明的，而几乎只是出于历史的偶然）。最后，人们猜想，很可能还有一种携带重力的"重子"（graviton），但现在还没有人发现过重子。

所有这些载体粒子在大尺度上的作用是，它们在宇宙中形成了很多"场"。重力的相互作用形成了重力场；电磁力形成了电磁场；而两种核力则共同形成了所谓的杨-米尔斯场（Yang-Mills field），这个场是以物理学家杨振宁和罗伯特·米尔斯（Robert Mills）的名字命名的。

我们可以将这些基本的力的特点总结到一张物理学家的购物单中：

- 重力：强度为 6×10^{-39}，作用范围无限，由重子（尚未发现，其质量可能为 0，自旋为 2）携带，来自重力场；
- 电磁力：强度为 10^{-2}，作用范围无限，由质子（质量 0，自旋为 1）携带，来自电磁场；
- 强核力：强度为 1，作用范围 10^{-15} 米，由胶子（质量 0，自旋为 1）携带，来自杨-米尔斯场的一个组成部分；
- 强核力：强度为 10^{-6}，作用范围 10^{-18} 米，由弱子（质量巨大，自旋为 1）携带，来自杨-米尔斯场的另一个组成部分；

你会觉得原来的36种基本粒子加上各种各样的胶子，虽然达到了60种之多，但这并不算是什么了不起的发展，不过是通过大量的对称严密地组织起来的夸克而已。它们都是同一主题的变奏，而不像夸克发现之前那样，物理学家必须应对各种杂乱无章的状况。

用夸克描述基本粒子的方法就是所谓的“标准模型”。这些描述和实验数据吻合得非常好。必须对其中一种粒子的团块进行调整，使其宜于观察，一旦你做到了这一点，其他的粒子团块也就井然有序了。这可不是什么循环论证。

夸克是被紧紧地捆在一起的，而且，你不可能看到单独的夸克，你能看到的只有两个或三个一组的夸克。但是，物理学家们已经间接地证实了夸克的存在，它们并不是强加在无秩序事物身上的数学命理学变种。谁要是认为宇宙在本质上是美的，那么，夸克的对称就会让你的想法更坚定。

*

根据量子色动力学，一个质子是由三个夸克组成的，其中包括两个上夸克和一个下夸克。如果你把夸克从质子中拿出来，并将其组合打乱，再把它们放回去，它们组成的仍然是一个质子。所以，夸克通过排列组合构成质子的法则应该是对称的。更有趣的是，事实证明，夸克类型改变的规律也是对称的。比如，你可以把一个上夸克换成一个下夸克，但这条规律仍然有效。

这也就意味着，实际的对称群并不是就是这三个夸克的六种排列组合，而是一个与之紧密相关的连续群SU（3），是基令发现的单群中的一个。SU（3）中的变换不会改变自然规律的方程，但是会改变

方程的解。比如，通过 SU（3），你可以把一个质子“旋转”成一个中子，只需把构成它的夸克颠倒一下就可以了，所以，两个上夸克变成了两个下夸克，一个下夸克变成了一个上夸克。费米子世界中就存在 SU（3）对称，它可以把一种费米子变成另一种费米子。

标准模型中还有两种非常重要的对称群。弱核力的规范对称 SU（2）可以把一个电子变成一个中微子。SU（2）群也是基令对称群表中的一个。我们的老朋友电磁场具有 U（1）对称，U（1）对称不是麦克斯韦方程中的洛伦茨对称，而是相变的规范对称。U（1）对称群不在基令的对称群表中，因为它不是 SU（1），但二者有着非常紧密的关系。

通过融合二者的规范对称，电弱理论统一了电磁力与弱核力。而标准模型又将强核力统一进来，它为所有的基本粒子提供了一条普遍适用的理论。这种方法非常直接：把三个规范对称群合成 SU（3）× SU（2）× U（1）就可以了。这种构成既简单又直接，只是不太优美，而且这让标准模型看起来像是用绳子和胶水粘连而成的。

假设你有一个高尔夫球、一个纽扣和一根牙签，高尔夫球具有球形对称 SO（3），纽扣具有圆形对称 SO（2），而牙签只具有单一的反射对称 O（1），你能找到一个同时包含这三种对称的统一体吗？当然可以，把它们三个放在一个纸袋子里就行了。现在，你可以通过 SO（3）旋转袋子里的高尔夫球，通过 SO（2）转动袋子里的纽扣，通过 O（1）把里面的牙签掉个头。这个纸袋子的对称群就是 SO（3）× SO（2）× O（1）。标准模式就是这样融合不同的对称群的，但标准模式中不会用到旋转，它用到的是量子力学中的“酉变换”。但是，它也有同样的缺点：它只是把三个系统堆到一起，并通过一种繁琐的方式把它们的对称合并起来。

还有一种更有趣的方法可以将这三个对称群合并到一个统一体中，而且比纸袋子漂亮得多。你也许会把牙签在高尔夫球上放平，再在它的一端把纽扣粘上。你甚至可以用一大把牙签作轮辐，让纽扣做轮毂，然后通过转动高尔夫球的顶部来转动这个轮子。如果你足够聪明，就可以让这个统一体具有多种对称，比如 K（9）（现实中并没有这个群，是我为了讨论而虚构出来的），单个的 SO（3），SO（2）和 O（1）很有可能有幸成为 K（9）的子群。你完全可能找到更巧妙的方法去合并高尔夫球、纽扣和牙签。

这也是物理学家正在摸索的关于统一模型的方法，他们还希望 K（9）就是基令对称群表中的一员，要么就和其中的对称群很相似，因为基令的群就是对称的基石。所以他们以 SU（5），O（10）和基令神秘的例外群 E_6 为基础，发明了一整套的"大统一理论"（Grand Unified Theory，或称 GUT）。

大统一理论和卡鲁查—克雷恩理论有着一样的缺点，它们都没能提出可验证的预言。但是后来，一个十分有趣的预言出现了。它是全新的，新得人们不敢相信它的正确性，可它竟然是可验证的。所有的大统一理论都预言质子可以通过"旋转"变成电子和中微子。所以质子是不稳定的，而且，长远看来，宇宙中所有事物都会衰变成射线。据推算，一个质子的平均寿命是 10^{29} 年，比宇宙的寿命长得多。但是个别的质子有时衰变得很快，如果有足够多的质子，就可以观察到它的衰变。

一箱水中的质子数量足够多，每年总会有一些会发衰变。到 1980 年代末，为了观察到质子的衰变，人们已经做了六次试验。其中最大的一个水箱装了三千吨极纯的水，但是没有一个人看到质子衰变。这就意味着，质子的平均寿命至少是 10^{32} 年，至少比大统一理论推测的

寿命长上一千倍。大统一理论根本就拿它没办法。回头想一下，如果人们观察到了质子的衰变，可就真有点尴尬了，因为大统一理论漏掉了一个很重要的东西，那就是重力。

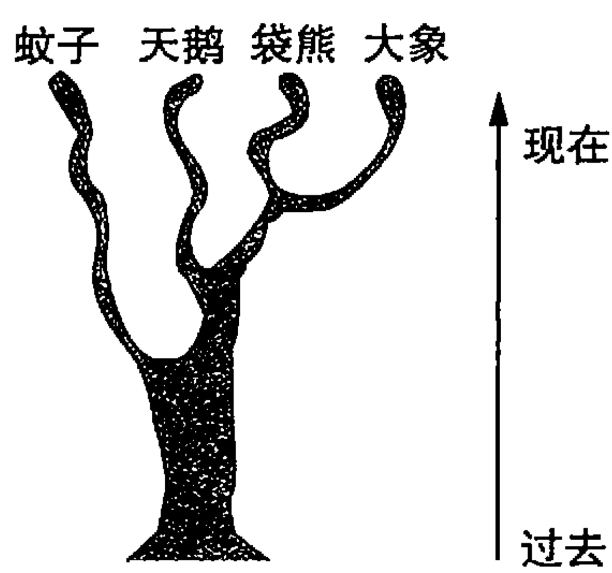
*

万有理论必须解释为什么只有四种基本的力，还必须解释它们为什么会具有如此惊人的形式。这有点像在一头大象、一只袋熊、一只天鹅和一只蚊子之间寻找家族相似性。

如果这四种力可以被当做一种力的四个方面，那它们就比较容易组织到一起了。在生物学中，人们已经取得了成功：大象、袋熊、天鹅和蚊子都是生物进化树的分支，它们由DNA统一起来，并通过DNA历史演变的程度区分开来。这四种动物都是从同一个祖先开始，经过了10亿年或20亿年的时间，一步一步进化而来的。

大象和袋熊的共同祖先比大象和天鹅的共同祖先出现要晚。所以，生物进化树上最新的分支就是这四种生物；往前追溯，大象和袋熊的共同祖先从它们与天鹅的共同祖先那里分离出来；再往前，这三种生物的祖先又是从它们和蚊子的共同祖先那里分离出来的。

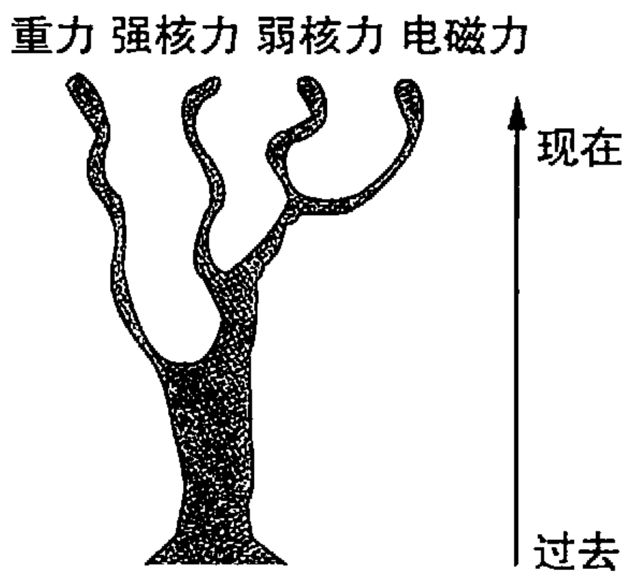
物种形成可以看做是一种对称破坏。单一的物种在物种内部是（近似于）对称的，每个袋熊都非常相似。出现两个不同的物种（袋熊和大象）时，你可以对袋熊进行排列组合，也可以对大象进行排列组合，但无论如何，你都不能把袋熊变成大象，也不能把大象变成袋熊。



四个物种随着时间分化

物理学家们对四种基本自然力的内在统一性的解释也与此类似。这里，扮演 DNA 在生物学中的角色的是宇宙温度，也就是它的能量级。尽管自然的内在法则永远不会发生变化，但是在能量级不同的情况下，它的表现形式也不同，就像水在温度低时是固体、温度适中时是液体、温度高时是气体一样。在极高的温度下，水分子就会破裂，变成等离子体，由分散的粒子组成；温度再高一些，这些粒子又会破裂，变成夸克—胶子等离子体。

宇宙 13 亿年前从大爆炸中产生时，温度非常高。最初，四种力的表现形式是一样的。但是随着宇宙的冷却，它的对称就被破换了，也就分散出了截然不同的单个的力。我们现在这个拥有四种力的宇宙是其优美源头的不完美投影，因为它是三次对称破坏的结果。



四种力是如何随着时间的推移分离出来的

政治记者

1972 年 6 月，美国大选正进行得如火如荼，水门综合大厦的一名保安发现一扇开着的门上被粘上了胶带。他取下了胶带，以为是工人不小心留下的，但是等他回来的时候，又有人把胶带粘上了。这让他警惕起来，赶紧打电话报警，结果，警察抓住了五个闯进民主党国家委员会办公室的男人。后来人们证实，这几个人和尼克松连任竞选委员会有关。

这个消息对选举本身没有产生太大影响，尼克松取得了压倒性的胜利。但是事情没有就此结束，水门事件的触手一层高似一层地探入了尼克松政府。《华盛顿邮报》的两个记者，鲍勃·伍德沃德（Bob Woodward）和卡尔·伯恩斯坦（Carl Berstein），对这个事件穷追不舍，并得到了“深喉”暗中提供的揭秘情报的支持。没有人知道深喉是谁，但是很明显是一位高级官员。2005 年，深喉的身份被揭开，他就是马克·费尔特（Mark Felt），联邦调查局的二号人物。

深喉泄露给媒体的信息是爆炸性的。1974 年 4 月，尼克松被迫辞

退了两名高级助手。后来人们发现尼克松在自己人的办公室里也装了窃听器，并且找到了记录着机密对话的录音带。这盘磁带的来源问题引发了一场激烈的法律论战，人们在磁带中找到了几段空白，很明显是有人故意对其进行了消磁处理。

人们普遍认为，努力掩盖入室犯罪和白宫之间的关系这件事比入室犯罪本身更严重。众议院已经开始启动了可能会导致总统弹劾的正式程序。弹劾是在美国参议院的主持下查证总统所犯的“重罪和轻罪”，如果被查明有罪，就必须离职。在弹劾和定罪在所难免时，尼克松辞职了。

尼克松竞选中的对头是参议员乔治·麦克格文（George McGovern）。他在南达科他州的苏福尔斯宣布代表民主党参选时有过一段预言性的评论：

> 今天，我们的公民不再感到他们可以和其他公民一起构筑自己的生活了。取而代之的是对我们的领导人的诚实性的不信任，以及对他表面上那一套东西的不信任。在美国的政治词汇表中，出现了一个令人痛苦的新词组，“信用裂隙”（credibility gap），也就是花言巧语与事实真相之间的裂隙。坦白来讲，这意味着人民已经不再相信他们的领导人告诉他们的东西了。

在麦克格文阵营的一些次要成员中，有一个准政治记者，如果麦克格文当选，他也就跟着飞黄腾达了。如果历史真是如此，那政治会很走运，而基础物理和高等数学却会很倒霉。在2004年的真实历史中，这个记者成了《时代》杂志选出的一百位全球最有影响力的人物中的一个，但却不是因为他的新闻报道。

他的入选是由于他对数学物理学作出的重大突破性贡献。他对很多世界上最具原创性的数学都有着重大贡献，并因此获得了菲尔茨奖，这是数学领域的至高荣誉，完全可以与诺贝尔奖相媲美。但是，他并不是一个数学家。他是当今世界上顶尖级的理论物理学家之一，获得过美国国家科学奖章，但是他的第一学位却是历史学。他是查尔斯·西蒙尼（Charles Simonyi）讲座的教授，供职于爱因斯坦工作过的普林斯顿高等研究院，他的名字叫爱德华·威滕（Edward Witten）。

威滕和可怜的狄拉克不同，而是和德国的量子理论家一样，是在知识分子家庭中长大的。他的父亲路易斯·威滕（Louis Witten）也是一个物理学家，研究广义相对论和重力。爱德华生于马里兰州的巴尔的摩，第一学位是在布兰戴斯大学（Brandies University）取得的。尼克松连任后，他在普林斯顿大学取得了博士学位，并在美国多所大学进行研究和教学工作。1987 年，他受雇于高等研究院，那儿的所有职位都只负责纯粹的研究，他现在还在那儿工作。

威滕的物理学研究是从量子场理论开始的，而量子场理论就是人们努力调和量子理论和相对论的第一个成果。在量子场理论中，运动的相对性也被考虑了进来，但只限于平直时空中的运动（由于重力必须在弯曲时空中，所以这里不涉及重力）。1998 年，在一次吉布斯讲座（Gibbs lecture）上，威滕说量子场理论“包含了目前人们所知的、除了重力定律之外的绝大多数物理定律。在它向前发展的 70 年中，有很多的里程碑，从反物质理论……到对原子的更精确描述……再到粒子物理的标准模型”。他指出，物理学家已大大发展了这一理论，但它一直缺乏数学的严密性，所以对数学没产生多大影响。

他说，弥补这一缺陷的时机已经成熟。理论数学中的几个主要领域正是伪装了的量子场理论。威滕自己的贡献，也就是对拓扑量子场

论的发现及分析，可以直接对为与物理学没什么关系的目的而发明的数学思想进行解释。这其中就包括英国数学家西蒙·唐纳森（Simon Donaldson）划时代性的发现——只有四维空间支持诸多不同的“可微分结构”（differentiable structures，使微积分得以实现的参数系统）。另一方面是纽结理论的最新突破，也就是著名的琼斯多项式；还有一种多维复杂曲面中的“镜像对称”（mirror symmetry）的发现；以及现代李群论中某些领域的发展。

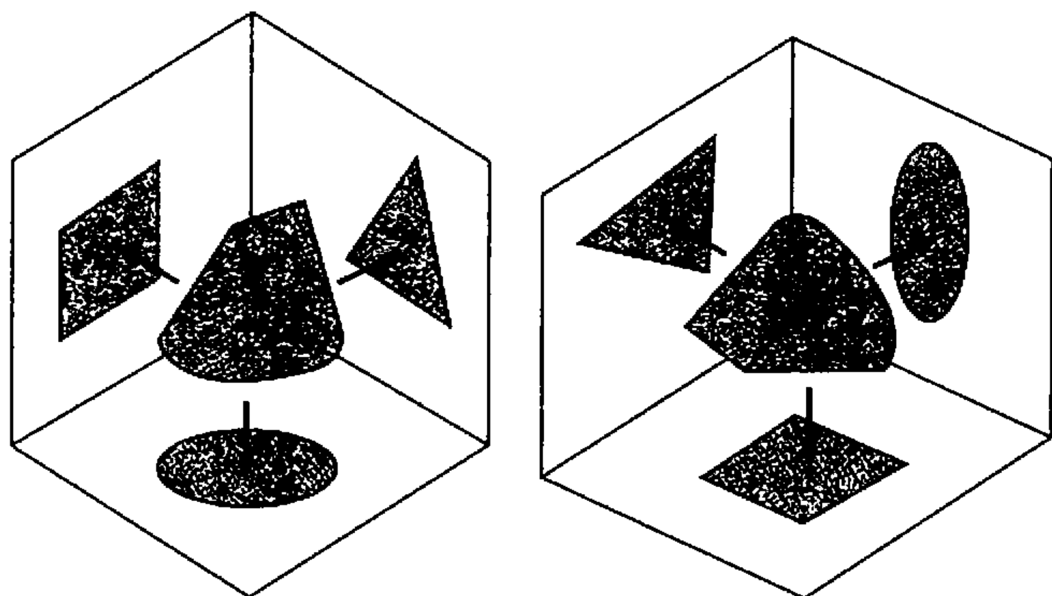
威滕作出了一个大胆的预测：21世纪的一个主要数学课题，将是量子场论中的数学思想与主流数学的合并：

> 这是一座巨大的山脉，它的绝大部分仍被云雾笼罩着。现在的数学理论发现的只是越过了云层的一个个山尖，而对这些山尖的研究也都是彼此孤立的……淹没在云雾中的是山脉的主体，它的床岩是量子场理论，而数学就是埋在山脉中的宝藏。

菲尔茨奖激励着威滕对这些宝藏的发掘，这其中就包括对“正质量猜想”（positive mass conjecture，在一个重力系统内，若局部的质量密度为正，那么，总质量也必然为正）更先进的新证明。这个猜想看起来不言自明，但是在量子物理领域，质量是一个非常微妙的概念。对这个人们寻找了很久的结果的证明，是1979年理查德·库恩（Richard Schoen）和丘成桐发表的，丘成桐还因此于1982年获得了菲尔茨奖。威滕先进的新证明研究了“超对称”，这是人们第一次把这个概念运用于重大的数学问题中。

*

我们可以通过一个古老的难题来理解超对称。我们需要一个瓶塞把瓶口塞上，而这个瓶子的开口的形状可以是圆的、方的，也可能是三角的。这些形状确实是存在的，而人们通常给出的答案是，瓶塞要像一个楔子，底部是圆的，一头是尖的。从下面看，瓶塞就像一个圆；从前面看，是一个正方形；从旁边看，是一个三角形。一个形体就可以同时具备这三种性质，因为三维的物体可以在不同的方向上形成若干个不同的“影子”或投射图。



超对称的原理　左图：一个瓶塞的三种形状；右图：瓶塞旋转后的效果。

现在让我们想像一下，一位平面世界的居民居住在我们图片中的“地板”上，他可以看到瓶塞在地板上的投影，却看不到另外的投影。有一天，他惊奇地发现，地面上的圆不知怎么变成了一个正方形。这怎么可能呢？正方形和圆形很明显是不对称的。

这是对称的，只是不是平面中的对称。但是在平面世界的居民转过脸去的时候，住在三维世界的人转动了瓶塞，所以它在地板上的投影变成了一个正方形，而旋转是三维空间中的对称转换。所以，更高维度中的对称有时可以解释低维度中的不对称现象。

超对称的情形与此极为相似，但不是从圆变成方，而是从费米子变成玻色子，这确实令人吃惊。这也就意味着，如果你可以对费米子进行推算，那么通过对称操作，你就可以不费吹灰之力地推测出玻色子的情况；反过来也是一样的。

我们在现实世界的对称中也可以看到这样的情况。你可以在镜子前拿几个球来玩戏法，镜子这边的你怎么做，镜子那一边的情况也会随之改变，镜子那边的戏法和你这边一定是相同的。如果镜子这边的你完成一整套戏法要用3.79秒的时间，那么你不用计就会知道，镜子那边完成相应戏法的时间也一定是3.79秒。镜子两边的情形构成了镜像对称，即，镜子一边发生的一切，都会在另一边得到反射。

超对称没有这么一目了然，但是其效应与此很相似。我们可以用超对称从一种粒子的情况推测另一种粒子的情况。这就好像是你可以进入宇宙中更高维度的空间，然后把费米子变成了玻色子。粒子是超对称地成对出现的：粒子会与自己的变形粒子形影不离，这个变形粒子就叫超粒子，电子和超电子成对，夸克与超夸克成对。然而由于历史的原因，与质子相匹配的不是超质子，而是光微子（photino）。有一个超粒子形成的“影子世界”（shadow world），与我们的日常世界发生着微弱的相互影响。

这种思想在数学上是十分优美的，但是这些人们预想出来的影子粒子质量太大了，根本无法在实验中进行观察。超对称很美，但很可能是不正确的。尽管不可能直接地对其进行证明，但间接的证明还是可能的，因为，主要的验证理论就贯穿于它的内在含义之中。

*

威滕孜孜不倦地研究着超对称，1984 年，他写了一篇名为“超对称与莫尔斯理论”（Supersymmetry and Morse Theory）。莫尔斯理论是拓扑学的一个分支领域，是以其创立者马斯顿·莫尔斯（Marston Morse）的名字命名的。英国最杰出的在世数学家米歇尔·阿提亚（Michael Atiyah）这样描述威滕的文章：“想要了解当今量子场理论的几何学家必读。文章还对莫尔斯的经典不等式作出了精彩的证明……这篇文章的真正目的是，通过无限维流形（infinite-dimensional manifold）为超对称量子场论找到理论根据。”接着，威滕又把这些技巧运用于前沿拓扑学和代数几何学中的一些热门问题上。

威滕不是数学家，这一点我在前面已经说得很清楚了，但这并不是说我认为他缺乏数学才能，而是恰恰相反。可以这样说，在我们这个星球上，没有人的数学才能可以超过威滕。但是，威滕身上的物理学直觉和数学才能一样卓然超群。

和数学家不同，物理学家们从来不羞于用物理学直觉掩盖数学逻辑上的缺陷。数学家们已经学会了用怀疑战胜盲信，不管有多么强大的证据支持，对他们来说，证明才是一切。威滕可以像数学家一样把自己的直觉和数学联系起来，这就是他的独特之处。阿提亚这样形容威滕：“他用数学形式解释物理学思想的能力是独一无二的。他能够出色地运用物理学的洞察力导出新颖而深刻的数学定理，一次又一次地让数学界大跌眼镜。”

但是这种非凡的直觉能力也产生了一些问题。威滕很多最重要的思想都是通过物理学原理和物理学分析得来的，而没有经过证明，而且，有一些思想至今仍没有得到证明。他并不是没有能力去证明，非

尔茨奖就可以表明他完全有这个能力，但是，他可以通过逻辑跳跃得出深刻而正确的数学结果，所以觉得根本没有必要去证明。

*

这里就有了一个很重要的问题，威滕美妙的数学和基础物理学有关系吗？难道说，他对美的追寻让他走进了数学的死胡同，和物理学中的真完全沾不上边？

到了 1980 年，物理学家们已经统一了四种自然力中的三个：电磁力、弱核力和强核力。但是大统一理论却把重力遗漏得一干二净。这种我们在日常生活中可以直接感受到的力，这种让我们双脚着地的力，竟然没有出现在这个综合体中，这确实令人难堪。

要写下把重力和量子力学理论合并起来的理论并不难，而且这个理论看起来还完全讲得通，但是，当人们尝试着去解这些由二者的合并而产生的方程时，这些方程就又变得讲不通了。其中最典型的问题是，本来应该代表合适的物理量的数字竟然有无限多个。在物理理论中，任何的无限性都是出现了某种错误的标志。这个无限性出现在激发普朗克将光量子化的辐射定律中。

有些物理学家开始相信，造成无限性的主要原因是，人们早已习惯于将粒子看做点。点是一个没有大小的位置，是一个数学虚构。而量子粒子已经将点模糊化了，但并不能将其斩草除根，下手必须再狠一些才行。1970 年代的一些先驱们甚至已经开始考虑，粒子也许是以某种极细的纤维绳（弦）为模型的东西。1980 年代，超对称理论也开始涉足这一问题，从而，超弦出现了。

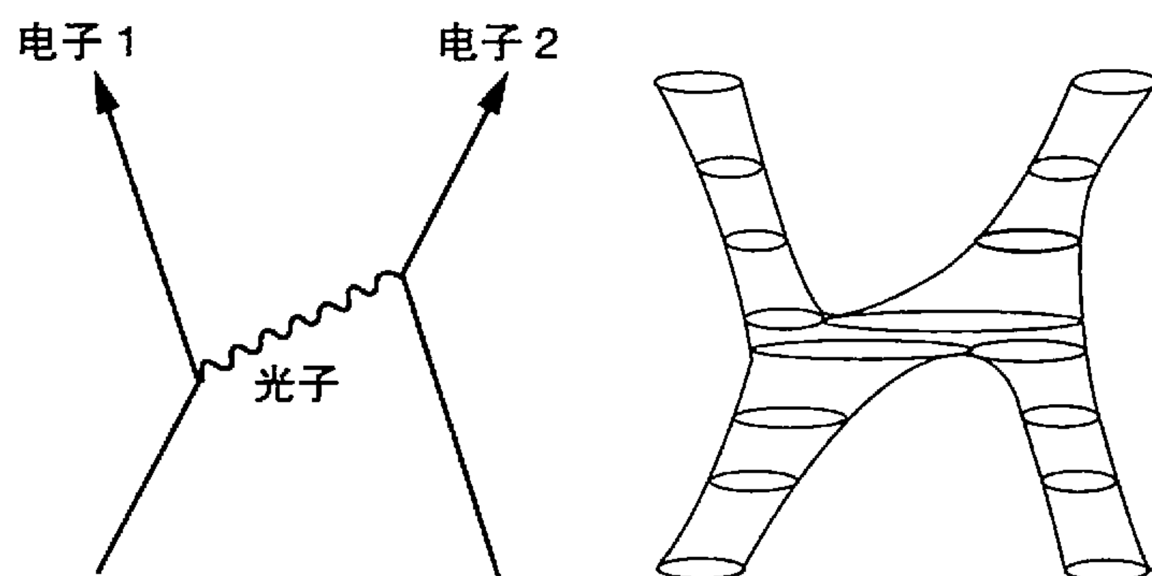
关于超弦理论完全可以写出一本书，而且有些人已经写出过这样

的书了，而我们只能浮光掠影地讲一讲而已。我在这里主要讲一下超弦理论的四个方面：合并相对论与量子物理学的方法、增加一个维度、以振动的形式存在于新增维度中的量子态以及这个新增维度中的对称，精确地说，是这一维度中各种各样的场。

我们必须从爱因斯坦用曲线表示粒子在时空中的轨道的思想开始，他将这种曲线称为世界线。本质上讲，这种曲线就是粒子在时空中的运动轨迹。在相对论中，世界线是平滑的曲线，这是由爱因斯坦场方程的形式决定的。它们没有分叉，因为在相对论中，所有系统的未来都是由过去（事实上指的就是现在）决定的。

量子场论中也有一个的概念，叫做费恩曼图（Feynman diagram），描述了粒子在图解式空间中的相互作用。例如，下图就是描述一个电子放射出的光子很快就被另一个电子捕获的费恩曼图。描述光子时，经常会用到弯曲测线。

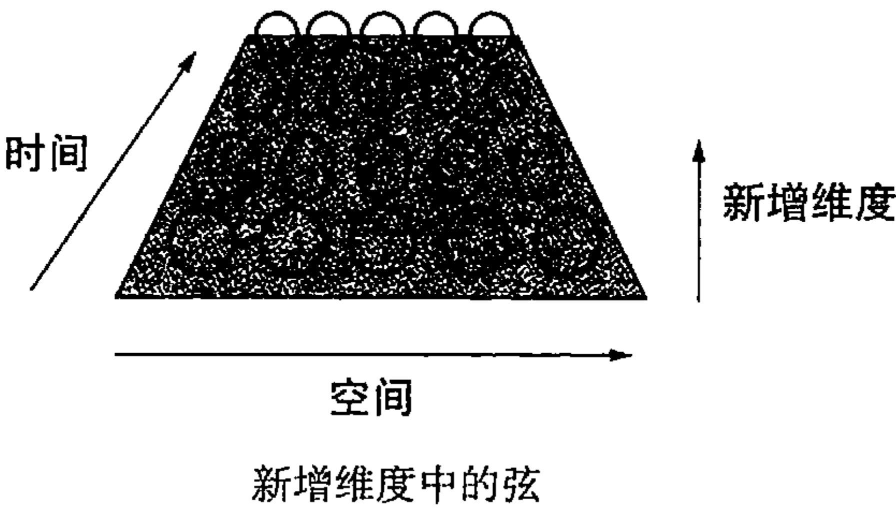
费恩曼图与相对论中的世界线有几分相似，但是有很多的尖角和分叉。南部阳一郎（Yoichiro Nambu）灵机一动地想到，如果不把粒子看做点，而是将其看做极小的线圈，那么，就能将费恩曼图转变成平滑的面，也就是世界面（worldsheet），如右图所示。世界面也可以看作一个修正过的时空中的世界线，只是新增了一个容纳这些线圈的维度。



（左）粒子相互作用的费恩曼图；（右）对应的世界面，被切成了一串薄片。

除了不是点，这些细线还有一个重要的特性，那就是它们会振动，几乎每种振动模式都对应着一种量子态。这就解释了为什么量子态总是以某个基数的整数倍的形式出现，比如，自旋数总是 1/2 的整数倍。线圈上的波的数量必须是整数。在一根小提琴的琴弦上，所有不同的振动模式都是根音及其高次谐波。这样，量子理论就变成了一种音乐，只是不在小提琴琴弦上演奏，而是在超弦上演奏。

南部的思想并不是无端地产生的，它源于加布里埃尔·维尼吉亚诺（Gabriel Veneziano）1968 年得出的一个伟大公式。用这个公式表示费恩曼图表示的同一个物理过程时，有着明显的区别，如果没能考虑到这种区别，就会在量子场论的运算中出现重大错误。南部阳一郎发现，当费恩曼图中布满了管道时，不同的图形就组成了有着相同拓扑结构的网络。也就是说，这些网络之间可以相互转变。这样，维尼吉亚诺的公式就与管道的拓扑性质联系起来了。



新增维度中的弦

因此，这就意味着具有不同的量子数（比如电荷数）的量子粒子，可能具有光滑时空（smooth space-time）的拓扑特征。数学家们已经发现了基本拓扑性质（比如说一个面上的洞的数量）的主要特点，至于细节，就复杂得难以言喻了。弦理论是把关于现实世界的所有细节全部统一起来的第一次尝试。

*

弦理论不是作为一条通向万有理论的途径而出现的，而是人们为了解释一种现在被称为强子的粒子而提出的。原子核中有很多普通的粒子，比如质子和中子，另外还有一些很奇特的粒子。但是这一理论存在着一个漏洞：它预测了一种质量为 0，自旋数为 2 的粒子，但人们从来没有观察到过（至今仍然没有）。此外，它没有想到还有一种自旋数为 1/2 的粒子，其实，大部分的强子，包括质子和中子，自旋数都是 1/2。这就和仲夏的天气预报只预报了会下一英寸厚的冰雹，却对温度的高低只字不提一样，这个问题没有引起物理学家们的注意。1974 年，量子色动力学取得了长足的进展，解释了所有的强子，甚至还预言了一种新的强子的存在，也就是欧米伽一负粒子，而弦理论却仿佛走到了尽头。

但是，就在这时候，约翰·施瓦茨（John Schwarz）和乔尔·谢尔克（Joel Scherk）发现，弦理论中不大容得下质量为 0，自旋数为 2 的粒子，而这种粒子很可能是具有长效重力的，也就是说，他们认为这一假想的粒子是带有重力的。弦理论会变成一种关于重力而非关于强子的理论吗？如果真是这样的话，那它就很可能成为一种万有理论，至少是一种关于很多东西的理论，因为还有很多的粒子不是强子。

这时，超对称派上了用场，因为它可以把费米子变成玻色子。强子既包括了费米子，也包括了玻色子，但还有一些粒子不是强子，比如电子。如果超对称可以与弦理论合并起来，那么，很多新粒子就会自动地随着超对称理论投奔到弦理论的旗下，因为超对称理论将会成为弦理论的一部分。

皮埃尔·雷蒙德（Pierre Ramond）、安德烈·奈沃（André Neveu）

和施瓦茨将这一合并起来的理论发展成了超弦理论。这种理论包含了自旋数为 −1/2 的粒子，并且摒除了一般弦理论中的一个令人头疼的特点，同时也排除了超光速粒子。现在，人们把这种粒子的存在视为其不稳定性的证据，并把它排除在外。

1980 年以来，英国理论物理学家米歇尔·格林（Micheal Green）通过李群和拓扑学技巧，研究出了很多关于超弦理论的数学内容。人们很快就发现，无论其物理学意涵是什么，超弦理论都具有令人惊异的数学美。但物理学仍然是桀骜不驯的：1983 年，刘易斯·阿瓦莱兹−高密（Luis Alvarez-Gaume）和威滕发现了一个潜藏在弦理论（包括超弦理论和更老的量子场论）中的暗礁。也就是说，这些理论中都存在“异常”现象。在把一个经典物理学系统转变成与其相似的量子系统时，就可能会出现这种异常。

施瓦茨和格林偶然地发现，只有当时空是 26 维（这一理论的第一个版本名为“玻色弦论”[bosonic string theory] 中）或是 10 维（在修正版中）的时候，这种异常才不会出现。为什么会出现这种情况呢？在他们对玻色弦论的运算中，造成异常的数学项是以 $d-26$ 为乘数的，其中 d 代表时空的维数。与此相似，在修正后的版本中，这个系数变成了 $d-10$。时间一直是一维的，但是空间有时会增加 6 个或 22 个维度。施瓦茨是这样解释的：

> 1984 年，我和米歇尔·格林就超弦理论作了一次运算，想看看这种异常到底会不会发生。我们的发现让自己大吃一惊。我们发现，一般情况下，确实会出现导致理论失效的异常现象。于是我们就可以自由地选择粒子的对称结构进行运算了，我们想先借此将这一理论明确下来。事实上，这样的对称结构有无数种可

能，但是，只有一种把这种异常从公式中神奇地抹去了，而其他的对称结构却不行。因此，在这无限多种可能性中，我们只找到了一个，它具有始终符合超弦理论的潜质。

如果你打算置10和26这两个古怪的数字于不顾，这一发现就会非常有意思。这说明，时空具有一定的维数是有其数学上的原因的。你会因这个数不是4而感到沮丧，但这还只是个开始。物理学家们一直都想搞清楚时空为什么有着这样的维数，过去人们会说："哦，它有多少个维度都有可能，但是在我们的宇宙中，时空是四维的，"但是现在，对于这一问题的答案已经不再局限于这样简单的一句话了。

也许其他的理论会得出一个四维的时空。这也许曾经是一个理想的时空，但是现在这种观念已经无效了，那些新增的维度已经在人们的头脑中挥之不去了。也许这些维度就在它在的地方。这是卡鲁查的一个老思想：也许存在着一个我们观察不到的时空维度。如果是这样，那么弦就占据着这个一维的线圈，但是这些线圈却必须在更高的维度中震动。粒子的量子数，比如电荷数和粲数，是由线圈震动的形式决定的。

一个很基本的问题是，这个隐秘的维度是什么样子的？时空是什么形状的？

起初，物理学家们希望这些新维度会形成一个类似于六维的圆环体之类的简单的形状。但是1985年，菲利普·坎德拉斯（Philip Candelas）、盖利·霍洛维茨（Gary Horowitz）、安德鲁·斯特劳明格（Andrew Strominger）和威滕推论说，更合适的应该是一种被称为卡拉比－丘流体（Calabi-Yau manifold）的形体。这种形体有上万个，下图是比较典型的一个：



一个卡拉比－丘流体（概要图）

卡拉比－丘流体的优点是，10维时空中的超对称就内嵌于四维时空当中。

这是例外李群第一次在前沿物理学中发挥如此巨大的作用，而且从此一发不可收拾。1990年前后，出现了五种形式的超弦理论，而且每一个的时空维数都是10。这些理论分别被称为Type I，Type IIA和Type IIB，以及Type HO和Type HE这两个“杂化”形式。有趣的规范对称群出现了，比如，在Type I和HO中，我们会发现SO（32），也就是32维空间中的旋转群；在Type HE中，我们会发现以$E_8 \times E_8$的形式出现的李群E_8，两个E_8起作用的形式是完全不同的。

例外群G_2也在超弦理论的最新成果中露了脸，威滕把这个新成果叫做M理论。他说“M”代表的是魔力（magic）、神秘（mystery）和母体（matrix）。M理论提出了11维时空的说法，将五种10维超弦理论统一了起来，这也就是说，只要将M理论的某些常量调整到特定值，就可以得出那10种理论中的任何一个。在M理论中，卡拉比－丘流形被一种名为G_2流形的7维空间代替了，因为这种空间的对称和基令的例外李群G_2有着密切的关系。

*

这时，弦理论的发展遇到了一点阻力，问题并在在于人们确凿地证明了它是错的，而在于人们无法确凿地证明它就是对的。一些杰出的物理学家，尤其是实验物理学家，从来没有对超弦理论做过任何实验——主要是因为超弦理论根本没有给出什么可做的事情，既没有可供观察的新现象，也没有可供测量的量。

我并不是认定超弦理论就是宇宙的钥匙，但是我仍然认为这种批评有失公允。有人要弦理论家证明自己的清白，然而很多时候，这种批评都上升到了非要证明他们有罪的地步。为了发现思考物理世界的新方法，人们已经花费了很多时间和精力，而弦理论在技巧上又有着很大的难度。原则上讲，弦理论完全可以对我们的世界作出预言，而主要问题是，必须经过异常繁琐的运算。40 年前，人们对量子场论也有过同样的抱怨，但是最终，通过更好的计算机和更好的数学方法，运算总算是完成了，而且，它与实验结果的一致性在科学史上是前所未有的。

另外，很多这样的指责都源于人们对万有理论（人们希望超弦理论是一种万有理论）的蕲求，但吊诡的是，理论越是完善，越是不容易证明，这是万有理论的内在性质决定的。为了能够取得成功，它必须符合量子理论，无论将其运用于什么样的实验中，其结果都必须和量子理论是一致的；它还必须符合相对论，无论将其运用于什么样的实验中，其结果都必须和相对论是一致的。所以，万有理论必须经过所有试验的检验。要让万有理论作出与传统物理学截然不同的预言，这无异于要求一种东西得出的结果和所有已经得到理论描述的物理现象都一致，然而又要求它与这些现象是不同的。

当然，弦理论迟早都要作出新的预测，并经过观测的验证，从一门纯理论转变成真正的物理学，必须符合所有我们已知的东西这一要求（至少到现在为止）并没有使预测变得不可能，只是说明了作出预测为什么会如此困难。已经有人提出了--些关键性试验的尝试性构想。比如，最近对银河宽度的观测表明，宇宙不只是在膨胀，而且是膨胀得越来越快了。超弦理论对此作出了一个简单的解释：重力泄露到了其他维度。但是，这一现象也可以用其他的途径进行解释。但是很明显，如果所有的理论家都不再研究超弦理论，那我们也就永远不会知道它是不是正确的。要找到一个具有决定性意义的实验需要花费大量的时间和精力，但还有一个前提，那就是这个实验确实存在。

*

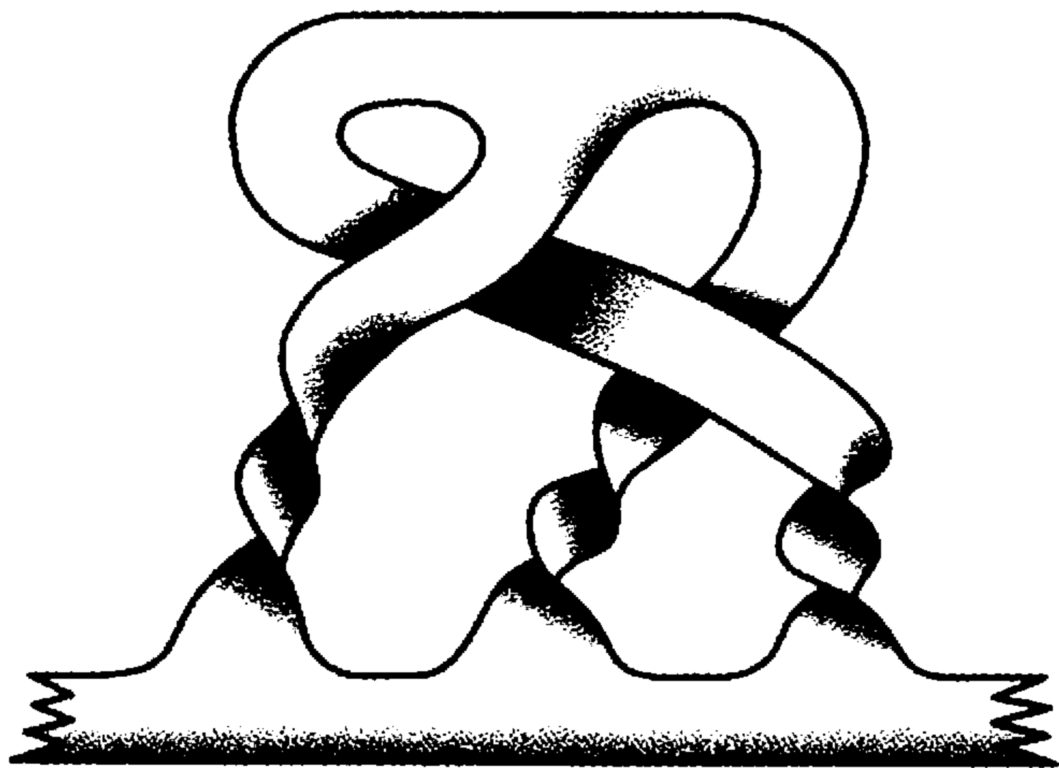
我并不是说超弦理论是统一量子理论和相对论的惟一途径，还有很多与之竞争的理论，但都同样缺乏实验支持。

其中有一种思想叫做“非交换几何”（noncommutative geometry），是法国数学家阿兰·孔涅（Alain Connes）创造的，建立在时空几何这一新概念的基础上。很多统一都是从时空是爱因斯坦的相对论模式的延伸这一思想开始的，然后想办法把亚原子物理中的基本粒子置入其中。孔涅则与此相反。他是从一种被称为非交换空间的数学结构开始的，这种数学结构包含了标准模式中出现的所有对称群，然后推导出类似于相对论的性质。关于这种空间的数学可以回溯到汉密尔顿和他的非交换四元数，只是已经经过了普遍化，也经过了修饰。但是孔涅这一理论同样是牢牢地建基于李群理论之上的。

另一个耐人寻味的思想是“圈量子重力”（loop quantum gravity）。

1980年代，物理学家阿贝·阿希提卡（Abhay Ashtekar）发现，爱因斯坦方程可以以量子理论的方式把空间看成是“颗粒状的”。这一思想是李·施莫林（Lee Smolin）和卡洛·罗威利（Carlo Rovelli）提出的，它导出了一种中世纪链甲式的空间，由 10^{-35} 米宽的团块组成，这些团块由纽带连接着。他们发现，这些纽带交织成辫（braid）或打成结的时候，这个链甲的细部结构就会非常复杂。但是，他们也弄不明白这种可能性到底意味着什么。

2004年，比尔森–汤普森（Sundance Bilson-Thompson）发现其中某些辫再现的正是夸克的组合规律。这些辫的拓扑结构重新对夸克的电荷进行了解释，其组合是通过“编织”这种简单的几何途径实现的。虽然这种思想现在仍处于起步阶段，但是它已经再现了标准模型中的绝大部分粒子。这是一大批对物质（这里指的是粒子）的设想中最新的一个，这些理论都认为物质只是空间中的“奇点”（singularity），也就是诸如结点、局域波或其他一些不平滑、不规则的空间结构。如果比尔森·汤普森是正确的，那么，物质就不过是扭曲了的时空而已。



用辫表示的电子

多年以来，数学家们一直在研究辫的拓扑结构，他们发现辫本身就形成了一个群，也就是辫群（braid group）。如果两个辫是通过两端连结起来的，就会用到“乘法”——我们在讨论鲁菲尼解决五次方程的方法时曾讲到如何把不同的排列联结起来，辫群的联结与此很相似。这一次，物理学同样建立在已经存在的数学发现的基础上，这些发现是数学“为着自己的目的”而发现的，仅仅是因为它们看起来很有趣。这一次，起关键作用的要素又是对称。

*

超弦理论最近面临的最大困境是自己过于富有了，之前的问题是没能作出什么预测，现在的问题则是作出了太多的预测。“真空能”（vacuum energy，没有物质的空间中的能）可以是任何东西，而它到底是什么则取决于弦在空间的额外维度中的缠绕方式。这些缠绕方式的种类数量极大，大约是 10^{500}，缠绕的方式不同，真空能的数值也就不同。

真空能的数值非常非常小，大约是 10^{-120}，但却不会是 0。

根据人们对“微调”（fine-tuning）的一般看法，这个数值恰好适合生命的存在。数值大于 10^{-118} 时，时空就会发生局部的爆炸；而数值小于 10^{-120} 时，时空就会陷入宇宙紧缩，最终会完全消失。所以，留给生命的“机会窗”是非常小的，而我们的宇宙能够幸存，真是一个奇迹。

“弱人择原理”（weak anthropic principle）指出，如果我们的宇宙不是以现在的构成方式构成的，我们也就不会在这儿了，问题是，为什么有这么个“这儿”让我们“在”呢？“强人择原理”（strong

anthropic principle）认为，我们之所以在这儿，是因为宇宙就是专为生命而设的——简直是神秘主义的谵妄。如果真空能与现存的状况有明显的不同，到底会出现什么样的情形？这一点尚无人能确切地说明。我们只知道有些东西会出问题，但是我们并不知道什么东西会变得“适合”（生命的存在），因为关于“微调”的绝大多数观点都是臆造出来的。

2000年，利用弦理论和真空能的10^{500}个可能值，拉斐尔·布索（Raphael Bousso）和约瑟夫·普金斯基（Joseph Polchinski）提出了另外一种观点。尽管10^{-120}非常小，而真空能量级的差值却还要小得多，大约只有10^{-500}。因此，很多弦理论都把真空能置于“适合”的范围内。然而要随机选出一个真空能量级，并要其适于生命的存在，这种机率还是极小的，但布索和普金斯基认为这并不要紧。事实上，“合适的”真空能量级的出现是必然的。这也可以说成是，宇宙也在考察所有可能存在的弦理论，并一直遵从着一个既定的理论，直到有一天，它达到了某种程度，就会通过量子力的方式“隧穿”至另一种弦理论。如果你能等上足够长的时间，那么，在某些阶段，宇宙就会获得正好适合生命存在的真空能量级范围。

2006年，保罗·斯泰恩哈特（Paul Steinhardt）和尼尔·图罗克（Neil Turok）提出了关于“隧穿”理论的另一种说法：周期性的宇宙在大爆炸中膨胀，又在宇宙紧缩中收缩，每过一兆年，就会重复一次。在他们的模型中，真空能在一连串的循环过程中会逐渐变小，因此，宇宙的真空能最终会变得非常小，但不会是零。

在另一个模型中，一个真空能足够低的宇宙会持续存在很长时间。这样，这个宇宙的情况就慢慢开始适合生命的出现，生命也从而获得了大量的时间，以便进化出智力，进而思考自己为什么会在这儿。

胡思乱想的数学家

鹅群是“欢叫”（gaggle）*，狮子是“骄傲”（pride），鸣禽是“妩媚”（charm），云雀是“嬉闹”（exaltation）……而数学家群体又如何命名呢？辉煌？有点太自鸣得意了。高深莫测？太浮泛了。当数学家们聚在一起时，通过对他们的“群体”行为的观察，我找到了一个最接近其真实情况的词——“胡思乱想”（muddle）。

其中一群“胡思乱想”发明了数学中最古怪的结构，并揭示了隐藏在其费解的表象后面的和谐。这个漫不经心地作出的发现，已经开始渗透到理论物理学中，而且很可能就是打开超弦理论的一些惊人性质的钥匙。

关于超弦的数学太新了，以至于其中还有很多未被发明出来。但具有讽刺意味的是，数学家和物理学家们发现，虽然超弦理论属于物

---

* gaggle 本身还有鹅群的意思，pride 又有狮群的意思，作者这里是为这些动物找了个近乎专有的形容词兼名词，而且希望为数学家群体找到一个合适的代词，使其获得一种感性形象。

理学的前沿领域，但却和维多利亚时代过时的数学有着异常紧密的关系，这确实令人吃惊。由于这种数学过于老旧，甚至数学课中都很少讲到。这一代数发明被称为八元数，是继实数、复数、四元数之后的又一种数系结构。

八元数是 1843 年发现的，但却被另一个人发表于 1845 年，后来人们也把这个功劳给错了人，但是没关系，因为没有人会在意这件事。到 1900 年的时候，甚至在数学界，八元数都被人遗忘了。1925 年，由于维格纳和冯·诺依曼试图将其作为量子力学的基础，八元数又经过了一次短暂的复兴，但在尝试失败后，它再次回到了从前默默无闻的状态之中。1980 年代，它们作为超弦理论一个得力的工具再一次浮出了水面；1999 年，它们又变成了 10 维和 11 维超弦理论的关键要素。

八元数告诉我们，8 这个数字有着某种奇异的性质，这种奇异性在物理学的空间、时间和物质中表现得更为明显。一个维多利亚时代的奇思异想摇身一变，同时成为揭开数学和物理学前沿难题的钥匙。在人们对于空间不止四维的信条，以及如何将量子理论和重力理论统一起来的问题中，八元数更是至关重要。

*

八元数的故事是抽象代数中一个争讼不休的话题，也是美国数学家约翰·贝兹（John Baez）发表于 2001 年的一本出色的数学研究著作的主题。在这里，我引入了贝兹的很多观点，我会尽力为大家讲清这个数学和物理的交界面古怪而又优美的奇特性质。就像哈姆雷特父亲的幽灵只有一个飘忽的声音一样，很多数学进展也是不易察觉的。请

原谅，我偶尔会用到一些专业术语，但也不要过于担心，有时候，为了便于研究一些主要人物，我才会用到这些术语。

我们最好先来做一下回顾。数系一步步的扩展与对称问题是交织在一起的。首先，人们在 16 世纪发现（或者说发明）了复数，在复数中，−1 也有了一个平方根。在此之前，人们一直认为数是神赐的、惟一的、不可改变的，没有人想过要去发明新的数。但是 1550 年前后，通过写出负数的平方根，卡尔达诺和邦贝利做到了这一点。人们花了 400 年的时间才理清了它的意义，但光是说服数学家相信它非常有用因而不可忽视，就用掉了其中的 300 年时间。

到了 18 世纪的最初几年，卡尔达诺和邦贝利的奇思怪想终于被总结为一种新的数，一般用 i 来表示。复数看起来也许很奇怪，但是事实证明，它是理解数学物理学的神奇工具。热、光、声、振动、弹性、重力、磁、电，以及流体流动等问题在它面前都会迎刃而解，可惜只有在二维物理中它才有效。

可是我们的宇宙有三个空间维度——至少到现在为止，我们一直都是这样认为的。既然复数的二维系统对于二维空间的物理这么有效，那么，有没有一个适用于真正的物理学的三维的数系呢？为了找到这样一个数系，汉密尔顿花了很多年的时间，却徒劳无功。1843 年 10 月 16 日，他突然冒出了一个想法：不要寻找三维数系，而要去找四维数系。他把自己的四元数方程刻在了布罗汉姆桥上。

*

汉密尔顿有一位大学的老同学名叫约翰 · 格雷夫斯，是一个代数迷。起初使汉密尔顿对数系的扩展产生兴趣的可能就是格雷夫斯。汉

密尔顿在把自己的四元数方程写在桥上后，给自己这位老伙计写了一封长信。格雷夫斯大惑不解，也不知道把一种新的乘法规则强加到四元数之上是否合理。“我对为了扩展数系而任意地创造一些虚数，并赋予它们超自然的性质，还没有明确的看法。”他在回信里写道。但是他也看到了这一思想的发展潜力，也对它到底能发展到哪一步充满期待：“如果你的炼金术可以搞到三磅重的金子，那干吗不继续下去呢？”

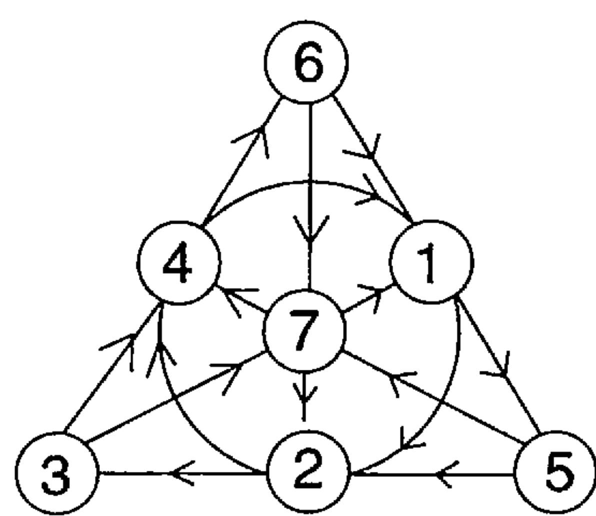
这是一个很好的问题，而且格雷夫斯准备自己来回答它。两个月后，他回信说自己发现了一个八维的数系，并把它叫做“octaves”(八元数)。他同时还发现了一个了不起的公式，包含了八个平方的和，我一会儿就会讲到。他还试图界定一个 16 维数系，但是由于碰上了他所说的“出乎意料的障碍”，所以没能完成。汉密尔顿说自己要把格雷夫斯的发现推广出去，但他一直忙于四元数的研究，根本就没空。后来，他发现了一个潜在的问题：八元数的乘法不符合结合律。也就是说，用两种方式——（ab）c 和 a（bc）——求三个八元数的积时，常常会得到不同的结果。在经过沉思之后，汉密尔顿准备忽略交换律，但最终却牺牲了有可能变成巨大掣肘的结合律。

格雷夫斯相当倒霉。在他发表八元数之前，凯莱也独力作出了相同的发现，并于 1845 年在一篇关于椭圆函数的蹩脚论文中将其发表了——还得寸进尺地将其收入了自己的作品集。凯莱把自己的数系叫做“octonions”。

格雷夫斯被这件事弄得闷闷不乐，更倒霉的是，不久就会刊发自己的论文的刊物正是凯莱之前发表自己的发现的刊物。所以，格雷夫斯为自己的文章加了一个注释，说明自己两年前就已经作出了这一发现。汉密尔顿也发表了简短的声明来支持他，说自己的朋友应该获得

优先权。虽然发现纪录得到了修改，但是八元数很快就被冠以“凯莱数”（Cayley numbers）的名字，这个名字现在还被人们广泛使用着。很多数学家都使用凯莱的术语，把这个数系叫做“octonions”，但却将其归功于格雷夫斯。这个名字从某些方面说比“cotaves”更好，因为它和“quaternions”很相似。

八元数代数可以用著名的法诺平面（Fano plane）来表示。这是一个由七个点和七条线组成的有限几何图形，每三个点由一条线连接，其大致形状如下：



法诺平面，由七个点和七条线组成的几何图形

为了能在该平面上画出来，其中有一条线必须弯成一个圆，但是不会产生什么问题。在这个几何图形中，任意两个点都由一条线连接，而任意两条线都会相交于一个点，图中没有平形线。当初发明法诺平面时，与八元数没什么关系，但后来人们发现，它和八元数的乘法规则正相符合。

八元数有八个单位：普通数字 1，以及另外七个分别名为 e_1，e_2，e_3，e_4，e_5，e_6，e_7 的数。这几个数字的平方都是 −1。这些单位的乘法规则是由图表决定的。如果你想将 e_3 和 e_7 相乘，可以先在图表中找到 3 和 7 这两个点，然后找到连接它们的线，在这条线上，还有第三个点，这个点就是 1。沿着箭头的方向，依次是 3，7，1，因此

$e_3e_7=e_1$。如果运算是逆箭头方向的，就要加一个负号：$e_7e_3=-e_1$。每对可能的单位都可以这样进行运算，这就是八元数算术。（加法和减法都很简单，而除法又是由乘法决定的。）

格雷夫斯和凯莱都不知道八元数和有限几何的这种联系，所以他们就不得不写出八元数的乘法表，法诺平面是后来才发现的。

多年以来，八元数都只是一个微不足道的怪东西而已。和四元数不同，它们既没有得到几何学的解释，也没有在科学中得到应用。甚至在理论数学中，八元数也没有发挥什么作用，毫无疑问，它只能变得默默无闻了。但是，在人们意识到八元数就是数学中很多著名的古怪代数结构的根源时，一切都改变了。它们解释了基令的五个例外李群——G_2，G_4，G_6，G_7 和 G_8——的真正来源；最大的例外李群 E_8 在构成 10 维弦理论的基础的对称群中出现了两次，而 10 维弦理论有很多让人惊喜的性质，很多物理学家都认为它很可能就是大家梦寐以求的万有理论。

如果我们同意狄拉克的观点，认为宇宙是按照数学的方式构成的，那么我们就可以说，由于八元数是存在的，所以 E_8 是存在的，而由于 E_8 是存在的，所以万有理论也是存在的。这就敞开了一种哲学可能性：我们认为宇宙的结构非常特别，而它之所以具有这种独具一格的结构就是因为它和一种独一无二的数学密切相关，这种数学就是八元数。

美即是真，真即是美。毕达哥拉斯主义者和柏拉图主义者都很乐意看到数学模型在世界的构建中起到关键性作用的证据。八元数具有一种惊人而离奇的数学美，正因为如此，狄拉克才认为 10 维弦理论是正确的；或者说，即使被证明是错误的，它仍然比那些正确的东西有趣得多。但是我们知道，美的理论未必真，而在对弦理论作出定论

之前，这种可能性都只能是纯粹的假设而已。

无论八元数在物理学中的重要性如何，关于八元数的种种思想都是数学中的纯金。

*

八元数和例外李群之间的关系，只是四元数的各种泛化形式与当今前沿物理学之间诸多奇特关系中的一种。我将在一定的深度上探讨一下其中一些关系，以便让你理解八元数有多么了不起。我将从数学中最古老的例外结构讲起，它就是平方和的公式。

这个公式很自然地来自于复数。每一个复数都有一个“范数”(norm)，也就是原点和它之间的距离的平方。根据毕达哥拉斯定理，可知 $x+iy$ 的范数是 x^2+y^2。复数的乘法法则是由韦塞尔、阿根、高斯和汉密尔顿奠定的，这一定理告诉我们，范数具有一种漂亮的特性。如果你将两个复数相乘，它们的范数也就会跟着相乘。用符号表示就是 $(x^2+y^2)(u^2+v^2)=(xv+yu)^2+(xu-yv)^2$。两个平方的和与另外两个平方的和相乘，总是等于两个平方的和。公元 650 年前后，印度数学家婆罗摩笈多就已经知道了这个事实，而费波那契是于 1200 年发现的。

早期数论家对两个数的平方和问题非常着迷，因为它们区分了两种截然不同的质数。如果一个奇数是两个平方的和，那么它一定可以写成 $4k+1$ 的形式（k 为整数），这一点很容易证明。还有一种奇数，其形式为 $4k+3$，不能表示成两个平方的和，即使两个平方之中有一个为零也不行。第一个这种例外数是 21。

费马作出了一个很漂亮的发现：这种例外数不可能是质数。他

还证明了相反的情况，每一个形式为 $4k+1$ 的质数都是两个平方的和。运用上面两个平方和相乘的公式，你会发现，当且仅当 $4k+3$ 中的所有质因子以偶次幂出现时，一个奇数才是两个平方的和。比如，$45=3^2+6^2$，是两个平方的和。它的质因式为 $3\times3\times5$，质因数 3 型为 $4k+3$，$k=0$，幂数是 2，而 2 是一个偶数。虽然另一个质因数 5 的幂数是奇数，但这个质数型为 $4k+1$（$k=1$），所以并不构成问题。

另一方面，例外数 21 等于 3×7，3 和 7 都是形为 $4k+3$ 的质数，在这里的幂数都是 1，而 1 是个奇数，因此 21 就不能表示为两个平方的和。这样的例外数有无限多个，而它们不能表示成平方和的原因都是一样的。

后来，拉格朗日运用相似的方法证明了每个正整数都是四个平方的和（0 也包括在内）。他的证明运用了欧拉 1750 年发表的巧妙公式，这个公式和前面我们讲到的公式相似，只不过是四个平方和，而不是两个。四个平方的和乘以四个平方的和，所得乘积也是四个平方的和。三个平方的和的公式很可能是不存在的，因为有几对数都是三个平方的和，而它们的乘积却不是。但是在 1818 年，德根发现了八个平方的和相乘的公式，这和格雷夫斯发现的八元数公式是同一个公式。格雷夫斯多么可怜——他关于八元数的发现是原创的，却被归功于他人；另一个发现，八平方公式，却又不算是首创。

另外，还有一个小公式，是关于一个平方的和相乘的，其结果还是一个平方。这个公式就是 $x^2y^2=(xy)^2$。这个公式只适用于实数，而二平方公式则适用于复数：其范数是“乘性的”（multiplicative）——一个乘积的范数也是不同范数的乘积。在这里，范数也是一个数和原点的距离的平方。一个数的负数和它的正数的范数是一样的。

四平方公式是怎么回事呢？它和四元数的功能是一样的。通过对毕达哥拉斯定理进行四维类推，我们可以得知，常见的四元数 $x+iy+jz+kw$ 的范数是 $x^2+y^2+z^2+w^2$，是四个平方的和。四元数范数也是乘性的，这也解释了拉格朗日的四平方公式。

也许不等我说你就想到了，德根的八平方公式也有着与八元数相似的解释。八元数范数也是乘性的。

讲到这里，我们会出现一种好奇的想法。我们有了四种类型的数系，一种比一种精密：实数、复数、四元数、八元数。它们的维数分别是 1，2，4 和 8。我们发现了相关的公式，一个平方和乘以一个平方和等于一个平方和：其中的平方数分别是 1，2，4 和 8。这些方程与数系是密切相关的，所以，最耐人寻味的还是数的模式。

1，2，4，8——下一个将是哪个数呢？

*

如果将这个模式推演下去，我们就会胸有成竹地期待找到一个 16 维的数系。事实上，这样一个数系可以通过一种正常的方法获得，这种方法就是凯莱-狄克森推展（Cayley-Dickson process）。对实数进行推展，就会得到复数；对复数进行推展，就会得到四元数；而对四元数进行推展，就会得到八元数。如果你继续往下对八元数进行推展，就得到十六元数（sedenion），一个 16 维的数系，然后就是 32 维代数和 64 维代数，如此等等，不一而足，每进行一步，数字就会增大一倍。

那么，这是否意味着有一个十六平方方程呢？

没有。因为十六元数的范数不是乘性的。平方和的乘法公式只在

平方数为 1，2，4，8 时有效。小数定律（law of small numbers）在这里又出现了：2 的幂的模式在这里止步不前了。

为什么呢？从根本上讲，这是因为凯莱－狄克森推展逐渐破坏了代数的法则。你每运用一次这种推展，得到的数系就与前面数系区别越大。如此以往，法则也跟着改变，原本井然有序的实数数系也就陷入了一片混乱之中。让我来详细地解释一下。

除了范数，这四个数系还有其他一些共同特征。最显著的特征是，它们都是“可除代数”，因此它们都是泛化了的实数，加法、减法和乘法这些代数运算对它们来说都是有效的。但是在这四种数系中，除法也是有效的。乘性范数的存在让它们变成了“赋范可除代数”(normed division algebra)。有一段时间，格雷夫斯曾认为从 4 维到 8 维的方法可以反复使用，可以导出 16 维、32 维、64 维的赋范可除代数，而且认为所有 2 的幂数都是如此。但是，他在十六元数上碰了一鼻子灰，并开始怀疑 16 维赋范可除代数存在的可能性。他是对的：现在我们知道，只存在四种赋范可除代数，它们的维数分别是 1，2，4，8，而且根本不存在与格雷夫斯的八平方公式和欧拉的四平方公式相似的十六平方公式。

为什么会这样呢？随着 2 的幂数的升高，新数系会一步步失去一部分结构。复数无法排列成一条线；四元数又违反了 $ab=ba$ 的法则，也就是“交换律”；八元数则违反了“交错律”(alternative law)$(ab)c=a(bc)$。而十六元数既不是可除代数，又没有乘性范数。

这不只是凯莱－狄克森的失败，而是一个更为根本的问题。1898 年，霍尔维茨证明了赋范可除代数只有原来的那四个老伙计。1930年，马克斯·佐恩（Max Zorn）证明这四种相同的代数是仅有的交错可除代数，只是几个不折不扣的例外。

这是专属于数学家的问题，就因为他们柏拉图主义的本能——爱这口。可是对于造福人类这个问题，仿佛只有实数和复数才重要，它们都有很大的实用性。四元数确实在一些深奥的问题上显示出了一些用处，可八元数在应用科学中的作用却微乎其微。它们就像理论数学的死胡同，一种只有在热衷于狂想的人身上才会发现的智力的谵妄。

*

数学史一再地告诉我们，不要因为一种又巧妙又美的思想没有明显的用途就将其弃之不顾。不幸的是，人们还是会不断地抛弃这些思想，常常就因为它们又巧妙又美。那些只考虑自己的人越“实际”，他们就越看不起从抽象问题中产生的数学观念，只在乎“对自己有用”，而不是站在真实世界（real-world）的立场上考虑问题。这些观念越美，他们就越是鄙视，仿佛美本身就是可耻的一样。

宣称这些观念无用，其实就是在跟命运打赌。一种被轻视的观念很可能会突然之间就变成万众瞩目的焦点，它不再是无用的，而是至关重要的。

这种例子不胜枚举。凯莱自己就说过他的矩阵一点用处都没有，但是今天，任何一个科学分支都缺不了它们。卡尔达诺说复数“微妙到了无用的地步”，但是每一个工程师和物理学家离了它都不行。1930 年代英国数学家中的翘楚戈德弗雷·哈罗德·哈迪（Godfrey Harold Hardy）说，幸好数论没什么用处，要不然人们就该把它用于战争了。现在人们通过数论将信息加密成密码，这成了保障网上商业活动安全的关键性技术，而它在军事中的作用就更为重要了。

八元数也是如此。它成了数学中的必修课，甚至在物理学中也是

一样。现在人们发现，它就是李群理论的核心，特别是那些对物理学有着重大帮助的李群，尤其是五个维数分别为 14，52，78，133 和 248 的例外李群，G_2，F_4，E_6，E_7 和 E_8。它们的存在是一个谜，一位数学家曾愤恨地说它们是上帝的暴行。

*

喜爱大自然的人们喜欢从不同的角度一遍遍地饱览美景圣地……半路上与瀑布不期而遇，从羊肠小道一旁的岩架上飞流直下，或站在海角上眺望蔚蓝的大海。与此类似，数学家也喜欢重温古老的课题，然后从一个全新的角度展开思考。基于对数学变革的理解，我们常常可以用全新的、深刻的方式来重新诠释过去的观念。这不仅仅是数学之旅的问题，傻呆呆地从不同的角度看着那些无用的东西，它还会帮我们找到新的、有效的方法，来解决各种老问题和新问题。在这方面，再没有比李群理论更鲜明、更富有教益的例子了。

我们前面讲过，基令把所有的单李群归纳为四个无限族，而其中有两个族可以归到另一个更大的族中，这就是偶维和奇维的正交群 SO（n）。另外两个族分别是特殊酉群 SU（n）和辛群 SP（$2n$）。

我们知道这些族都是同一主题的变奏。它们都是由满足一种特殊的代数条件的 $n \times n$ 矩阵构成的，这个代数条件就是它们都是“斜厄密”矩阵。它们之间的惟一区别就是，只有通过实数矩阵才能得到正交李群；只有通过复数矩阵才能得到酉李群；只有通过四元数矩阵才能得到辛李群。这些代数有无限个族，因为矩阵的大小是无限多样的。如果看到李群就是汉密尔顿力学的变体，可以用四元数的术语对其进行描述，你会觉得非常有趣，因为汉密尔顿力学是他的第一个发

现，而四元数正是汉密尔顿的最后一个发现。

用八元数描述矩阵时，也会出现令人惊奇的现象。不幸的是，由于缺乏结合性 (associativity)，你无法得到一个关于单李群的新的无限族。事实上，这也是一种“幸运”，因为我们知道这样的族根本就不存在。但是如果你对八元数进行了正确的操作，却会出乎意料地得到李代数。

1914 年，卡当问了一个很明显的问题，并给出了一个惊人的答案，八元数矩阵的课题也就因此而展开了。数学和物理学中有一条指导原则：如果你对某一研究对象产生了兴趣，首先要问的就是它的对称群是怎样的。实数的对称群没什么可讲的，它只具有“什么也不做”的恒等变换；复数的对称群是由恒等变换和镜像变换组成的，可以把 i 变换成 $-i$；四元数的对称群是 SU（2)，和三维空间中的旋转群 SO（3）很相像。

卡当问道，八元数的对称群是什么呢?

如果你是个卡当一样的人物，你也可以对这个问题作出回答。八元数的对称群就是最小的例外李群，人们将其称为 G2。八元数的八维数系有一个 14 维的对称群。例外赋范可除代数与第一个例外李群是直接相关的。

*

要进行下一步的讨论，我们还要用到另一种思想。我们必须回到文艺复兴时期——这次要讲到的是艺术家，而不是数学家。

在当时，艺术与数学的联系非常紧密，这不仅仅是说建筑和绘画。文艺复兴时期的画家们学会了如何运用几何透视，他们发现了使

纸上的图像看起来就像现实中的三维物体和景色一样的几何法则。就这样，他们发明了一种美轮美奂的新几何。

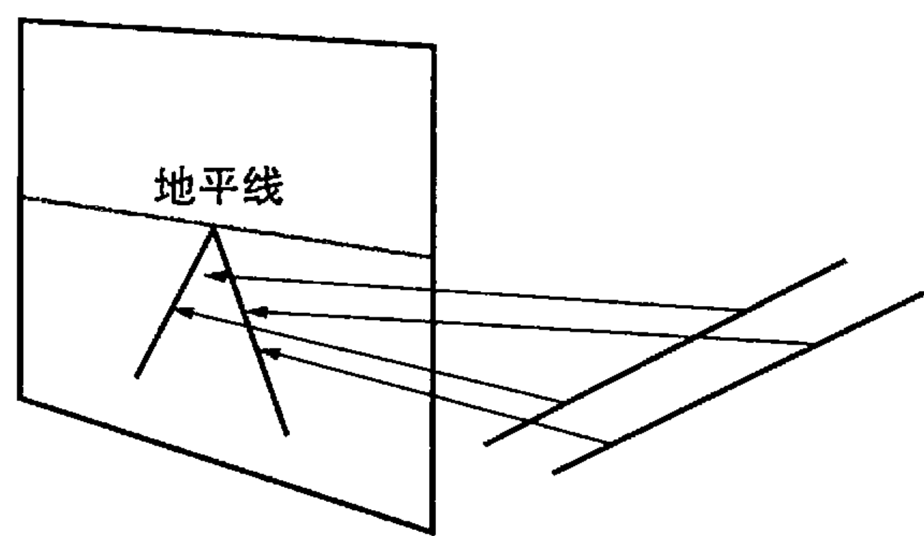
在此之前的艺术家的画作看起来不够真实。即使是乔托这样的画家想把画画得像照片一样，但仔细去分析的话，其比例也并不完全合理。1425 年，菲利波 · 布鲁内莱斯奇（Filippo Brunelleschi）发现了获得精确透视系统的数学方法，并把这种方法教给了其他的艺术家。1435 年，出现了第一本关于透视的著作，这就是莱奥尼 · 阿尔伯蒂（Leone Alberti）的《绘画》（*Della pittura*）。

这种方法在皮埃罗 · 德拉 · 弗朗切斯卡（Piero della Francesca）的创作中得到了完善，因为他同时也是一个高明的数学家。皮埃罗写了三本书来讲对称数学。这里我们还必须提到达 · 芬奇，他的《绘画专论》（*Trattato della Pittura*）是这样开头的："只有数学家才能读我这本书，"简直就是古希腊柏拉图学园大门上的标语"不懂几何者莫入"的遥远回声。

透视的本质就是"射影"（projection），通过（观念性地）画出人眼看到每个点，找出线与纸的焦点，将三维的景物在一个平面（纸）上描绘出来。射影的核心思想是使物体变形，而在欧几里得看来，这是不允许的，尤其是，射影可以把平行线变成相交线。

这种效应我们每天都会看到。当我们站在桥上看一条又长又直的铁路或高速公路时，它们会在远处消失，这些线会向中间合拢，并在地平线上相交。线与线之间的真实距离并没有改变，但是随着这些线向远端延伸，透视就会使观念中的距离缩小。在数学中，如果平面中的两条无限长的平行线获得了合适的射影，它们也会相交。但是它们的相交在平面中是不可想像的——它们不会在平面中相交。很明显，所有的线和平面都是向"地平线"延伸的。在平面上，地平线是无限

远的，但是，射影是把线从画面中穿过去。



射影如何使平行线相交

这种线被称为“无穷远线”（line at infinity）。和 −1 的平方根一样，它是虚构的，但是非常有用。关于这一现象的几何叫做射影几何（projective geometry），在克雷恩的爱尔兰根纲领中，这种几何讲的是，通过射影使景物的性质保持不变。所有艺术家用透视法作画时都通过地平线和“灭点”（vanishing point）使自己画中的图像看起来和真实景物一样，他们用到的都是射影几何。

射影平面几何是非常优美的。任意两点都可以通过惟一的直线连接起来，这和欧几里得几何是相同的；但是任意两条线都会相交，这就和欧几里得几何大相径庭了。欧几里得几何中的平行线在这里是不存在的。

如果你这时想起了法诺平面，那就对了，法诺平面就是一种有限射影几何。

*

从文艺复兴时期的透视到例外李群只是一小步。射影平面表明，阿尔伯蒂的方法暗含着一种新的几何。1936 年，一个后来成了建筑师

和工程师的军官杰拉德·德札尔格（Girard Desargues），发表了《尝试处理圆锥与平面相交情形的文稿》（*Proposed Draft of an Attempt to Deal with the Events of the Meeting of a Cone with a Plane*）。从名字看，这好像是一本研究圆锥曲线的著作，没错，但是他运用的不是传统的古希腊几何，德扎尔格用的射影法。因为欧几里得几何可以通过德扎尔格的坐标系（x，y，一对实数）变成代数，所以通过将 x 或 y 变得无限大（通过一种控制得非常巧妙的方法。其中会牵涉到三坐标的比例，以及 1÷0= 无限大），也可以将射影几何转变成代数。

你能够对实数进行怎样的操作，也就能对复数做同样的操作，这样，你就得到了复数射影几何。既然复数可以做到，那我们何不试试四元数和八元数呢？

这样做是有问题的——这种一目了然的方法失效了，因为在此过程中缺乏可换性（commutativity）。但是 1949 年，数学物理学家帕斯考尔·乔丹（Pascual Jordan）找到了一种颇具启发性的方法，成功地构造出了具有 16 个维度的八元数射影平面。1950 年，群理论家阿尔芒·博雷尔（Armand Borel）证明第二个例外李群 F_4 是八元数射影平面的一个对称群。八元数的摄影平面和复数的摄影平面很相似，但是必须用两把以八元数为刻度（而不是以实数为刻度）的八维 “尺子”来构造。

所以，现在已经有两个例外李群获得了八元数解释。那么，剩下的三个，E_6，E_7 和 E_8 呢？

*

八元数是上帝的暴行的观点在 1959 年已经流传很广了。而就在

这时，汉斯·弗雷登萨尔（Hans Freudenthal）和雅克·蒂茨（Jacques Tits）各自独立地发明了魔方，对 E_6，E_7 和 E_8 作出了解释。

魔方的行和列分别对应了四种赋范可除代数。任意给出两个赋范可除代数，然后运用数学技巧在魔方上找出对应的行和列，这时魔方所呈现的就是一个李群。其中一些群是一目了然的，比如，与所有行和所有列相对应的是三维空间中的旋转群 SO（3）。如果行和列都与四元数相对应，你就会得到 12 维空间中的旋转群 SO（12），数学家对这个群是相当熟悉的。但是你要找到八元数的行和列，就必须通过 F_4，E_6，E_7 和 E_8 这几个李群。被我们漏掉的 G_2 与八元数的关系也非常紧密——我们前面已经了解到，它就是八元数的对称群。

所以，现在人们普遍认为，由于大智大能的神允许八元数的存在，例外李群才得以存在。关于这一点，我们应该已经很明了了。就像爱因斯坦所说的，上帝难以捉摸，但他没有恶意。这五种例外李群全都是各种八元数几何的对称。

1956 年前后，前苏联几何学家博雷斯·罗森菲尔德（Boris Rosenfeld）猜想，E_6，E_7 和 E_8 这三个例外李群也是射影几何的对称群。这里不要你想八元数问题，但你必须理解下面这几种结构：

- E_6 结构：原八元数（biooctonion），由复数和八元数构成；
- E_7 结构：四分八元数（quateroctonion），由四元数和八元数构成；
- E_8 结构：八分八元数（octooctonion），由八元数与八元数自身构成。

这里遇到的惟一的小麻烦就是，没人知道如何通过数系的混合来

界定射影平面。但是已经有证据显示，这种思想是讲得通的。我们现在已经能够证明罗森菲尔德的猜想了，但是只有一种方法，那就是用群来构造射影平面。但这是不够的，因为这种思想必定会走上另外一条路，从射影平面走向群，但这还仅仅是个开头。事实上，E_6，E_7各自都获得了构造射影平面的不同方法，只有E_8仍然没有得到属于自己的方法。

*

正如基令所期望的，如果没有E_8，李群的故事就简单得多了，但也会乏味很多。我们世人是无权选择的：八元数和其他与我们如影随形的东西一样，就这么存在着。在某种我们尚不明了的意义上，宇宙是有赖于它而存在的。

八元数与生命、宇宙及万物的关系都是在弦论中出现的。弦理论的关键特征是，必须要有一个额外的维度来容纳弦的存在。原则上讲，这个额外的维度可能有很多的形状，最大的问题是要找到合适的形状。在旧式的量子理论中，关键的原则是对称，在弦理论中也是如此，所以，离了李群是万万不行的。万事万物的命运都系于关于对称的李群之上，例外李群更是尤其重要——它会在某种万一的时刻帮助物理世界正常地运转。

正是由于这一点，我们才回到了八元数问题上。

这里有一个关于八元数的作用的例子。1980年代，物理学家发现，3维、4维、6维和10维时空之间有一种相当可爱的关系。向量（直线长度）和旋量（保罗·狄拉克在电子自旋理论中首创的代数概念）与这些维度有着井然有序的联系。为什么呢？人们发现，赋范可除代

数的维数比空间维度大 2 时，向量—旋量关系是保持不变的。将 3，4，6 和 10 减去 2，得到的结果正是 1，2，4，8。

从数学的角度看，在 3 维，4 维，6 维和 10 维的弦理论中，每一个旋量用两个数表示成相关的赋范可除代数，而在其他的维度中不行，这为物理学带来了很多有益的影响。因此，我们现在就有了四种可供选择的弦理论：实数的、复数的、四元数的和八元数的。在这些备选的弦理论中也出现了这样的现象，当今人们认为最有可能与实在（reality）相对应的就是 10 维弦理论，它是由八元数所规定着的。如果这个 10 维弦理论真的符合了实在，那么，我们的宇宙就是以八元数的形式构建的。

这些奇异的“数”（它们甚至不能叫这个名字，因为它们不过是满足了某些代数法则而已）已经不止一次地起到如此重大的作用了。时下弦理论中的大热门 M 理论和 11 维时空有关。为了把 11 维时空变成我们熟悉的 4 维时空，我们必须扔掉 7 个维度，把它们紧紧卷起来，卷紧到无法察觉的程度。但是，你拿什么来应付 11 维的超引力呢？用八元数的对称群，也就是例外李群 G_2 就可以了。

它们再次露面了：不是作为维多利亚时代的古董，而是作为一种可能的万有理论的有力导引。这是个八元数的世界。

真与美的追寻者

济慈说得对吗？美即是真，真即是美吗？

二者紧密相连，也许是因为我们的思维对它们的反应是相似的吧。数学中可行的在物理学中并不一定可行，反之亦然。数学与物理学的关系深奥、微妙、令人费解。科学怎样揭示了自然界的鲜明“法则”，自然界又为什么喜欢用数学的语言言说，这是一个关于最高秩序的哲学谜题。

宇宙本然就是数学的吗？抑或它鲜明的数学特征只是人类的发明？还是说对于无限复杂的自然界，我们能够理解的最深的一面就是数学？

数学并不像人们曾经认为的，只是终极真理的某种空洞版本。如果说这本书讲了点什么，那就是，创造了数学的，是人。通过他们的成功和挫折，我们很容易验证这一点。加洛瓦和阿贝尔都死于21岁，我们怎能不为他们的骤逝感到震惊？一个是为人深爱却一直没有足够的钱结婚；另一个则天资颖慧又性情急躁，爱上别人却遭到拒绝，而

且很可能是因爱而死。当今的医学进步也许可以拯救阿贝尔，而且也很可能会帮助汉密尔顿保持清醒。

由于数学家也是人，也生活在普通人中间，所以数学上的新创造也是社会进程的一部分。然而无论数学还是科学，都不像社会相对主义者所言，只不过是社会进步的结果。此二者都有必须尊重的终极法则：数学中的法则是逻辑，科学中的法则是实验。但是，就算数学家们绞尽脑汁，想要用欧几里得几何学三等分一个角，也是枉费工夫；而无论物理学家多么希望牛顿的重力定律就是对宇宙的终极描述，水星轨道顶点处的运动都拒不服从这一定律。

这就是数学家们为什么都是死脑筋，沉迷于那些大多数人都不会留意的犄角旮旯。五次方程能不能用根式求解真的就那么重要吗?

历史对这一问题的回答是不容置疑的，当然重要。这些问题可能对日常生活来说并不重要，但是对于全人类来说是至关重要的。其重要性不在于我们可以解出五次方程，而在于弄清为什么我们无法找到那扇进入一个崭新的数学世界的神秘之门。如果加洛瓦及其后继者从未沉迷于五次方程是否可用根式求解的条件问题，人类发现群论的时间就会被大大延迟，也可能永远都不会发现。

也许，你在厨房里或在去上班的途中用不到群，但是，如果没有群，当今的科学就会大为失色，我们的生活也会与现在迥然有别。重要的不是大型客机、GPS 导航或手机的发明（这当然也是很重要的一部分)，而是理解自然的方式。没有人会想到一个迂腐的解方程问题会揭示物理世界的深层结构，但事实就是如此。

历史就是最好的明证。对深奥的数学问题的研究不会因其没有立竿见影的用处而受到贬斥。好的数学比黄金更有价值，而其最初被发现时却往往显得不合时宜，重要的是它将导向哪里。

*

令人吃惊的是，最好的数学往往导向难以预料的领域，而且事实证明，其中有很多对科学技术来说都至关重要，然而发明它们的最初目的却完全不是这样的。椭圆在古希腊只是圆锥体研究的一部分，经过开普勒，再到第谷·布拉赫（Tycho Brahe）对火星的观察，最终发展出了牛顿的万有引力理论。矩阵理论的发明者凯莱抱怨它毫无用处，而它却变成了统计学和经济学的基本工具，成了每一门科学的重要分支；八元数理论可能会催生一种万有理论。当然，如果剔除了相关的物理学内容，超弦理论可能只是一个漂亮的数学片段。如果是这样，当今量子理论中对对称的应用同样说明，群论提供了一种理解自然的深刻视角，而最初它可能只是为了回答一个纯数学问题。

为什么数学对其发明者完全没有想过的目的具有这么大的用处呢？

古希腊哲学家柏拉图说“上帝是几何化的”。伽利略也说过类似的话：“自然这部大书是用数学的语言写成的。”开普勒找到了行星运行轨道的数学模式。牛顿从其中一些模式得出了万有引力定律，而其余的则成了无用的玄想。

很多现代的科学家也谈论过数学思考的惊人力量。维格纳认为数学“不可用理智衡量的力量”（unreasonable effeetiveness）是理解自然的门径，这个短语出现在他写于1960年的一篇文章中。在文章的正文中，他说自己要阐明两个主要的观点：

> 第一点是，数学在自然科学中那近乎神奇的巨大作用很难用理性解释。第二点是，正是数学思想这种鬼斧神工的作用导致了

物理理论的独立性问题。

他还说：

数学语言对物理定律神奇的适用性是一个绝妙的礼物，而我们既无法理解也不配领受。对此，我们应该感激，并希望这种适用性在以后的研究中继续有效，甚至增强这种有效性，尽管这同时有可能扩大我们的困惑和学习面。

保罗·狄拉克认为，自然的法则应该不仅是数学的，还应该是美的。这就是说，真与美应该是一体两面的，数学的美引导着物理的真。他甚至说，与一个正确的理论相比，自己更喜欢一个美的理论，他认为美高于简洁："一个研究者在将自然的基本法则表述为数学形式时，应该致力于表达数学的美。与美相比，他应该始终把简洁看做次一级的东西……必须优先考虑是否有什么是与美相冲突的。"

有意思的是，狄拉克关于数学美的观点与大多数数学家都是迥然不同的。这种观点的确没有考虑到逻辑的精确性，他的著作中也存在着很多逻辑上的漏洞——最著名的例子就是他的点突函数（delta function）是自相矛盾的。但是，这个"函数"在他那里发挥了重要作用，最终，数学家们在思想的精确性方面对它进行了修正——被修正的正是他认为很美的地方。

尽管如此，狄拉克的传记作者黑格尔·科拉夫（Hegle Kragh）还是说："（狄拉克）所有的伟大发现都是（20 世纪 30 年代中期）以前作出的，而 1935 年之后，他几乎对物理学没有任何贡献。这与数学之美在后来这段时期控制了他的思考不无关系。"

的确是不无关系，但这样说也许并不正确。也许狄拉克后期明确表达了这一原则，但在此之前，他已经开始使用这一原则了。他最好的著作在数学上都是非常优美的，而且把优美看做判断自己思维成果的标准。他没有说数学上的美就是物理上的真，而是说数学上的美对物理上的真是“必要的”。但这并不是充分条件。很多美的理论一旦面对实验，就变得毫无意义。正如托马斯·赫胥黎所说：“科学是有组织的常识，在科学中，美丽的理论往往败给丑陋的事实。”

但是有很多证据表明，自然在本质上是美的。其研究连接起了群论和物理学的数学家赫尔曼·外尔说：“我在很多著作中都尝试着统一真与美，但有时不得不有所取舍，而我通常会选择美。”量子力学的奠基人海森堡给爱因斯坦写信说：

> 如果我说我要把简洁与优美当做真理的美学标准，你也许会反对，我可以毫不掩饰地承认，数学命题的简洁与优美深深吸引着我，这正是自然的馈赠。你一定也有同感：大自然在我们面前突然绽放出的关系，有一种近乎骇人的简洁和浑整。

接着，爱因斯坦发现，有那么多最基本的东西都是未知的——时间的本质、万物秩序的来源、宇宙的形状——因此我们必须提醒自己，要理解“终极”之物，还有太长的路要走。从有用性方面说，数学只给了我们逻辑的和暂时的真理。但无论怎么说，它都是我们前进的最佳方向。

*

纵观整个历史，有两大源头充实着数学。其一是自然世界，另一个较抽象的世界是逻辑思考。正是二者的融合赋予了数学解释宇宙的力量。狄拉克完美地理解了这种关系："数学家的游戏规则是自己发明的，物理学家的游戏规则是自然赋予的，但随着时间的流逝，越来越多的证据显示，那些数学家感兴趣的规则也正是自然所选定的。"理论数学与应用数学是相互补充的。它们并非背道而驰，而是两个相互联系的思维光谱的端点。

对称理论的故事证明，即使是对一个高端数学问题（"五次方程可解吗？"）的否定回答，也会催生深刻的、根本性的数学。重要的是，为什么答案是否定的。解释这一问题的方法同样适用于其他很多问题，其中就包括深奥的物理学问题。但我们的故事也证明，数学的健康发展有赖于物质世界通过对现实生活的改造为其注入活力。

数学的真正力量正来自于人类对形式（美）的认识与物质世界的融合。人类的认识需要在物质世界中进行检验（真），并将其作为取之不竭的灵感源泉。没有新的数学思想，我们就无法解决那些看似科学的问题。但是如果新思想本身走向了极端，也会变成毫无意义的游戏。科学要求数学不断进步，并且常常需要新的数学思想。

如果数学只是被这些需求所驱动，成为科学的奴隶，也就只能像奴隶那样工作了——昏昏沉沉，勉勉强强，磨磨蹭蹭。如果数学研究只顾及自身的要求，就会变成一个顽劣、自私的家伙——养尊处优，自我中心，而且满脑子都是自己的重要性。最好的数学应该平衡自身的要求和外在世界的要求。

数学的难以衡量的影响正来源于此。平衡的个性是通过经验学来

的，还要将学到的东西转换到新环境中去。现实世界催生伟大的数学，但是伟大的数学却可以超越自己的源头。

那个发现了二次方程解法的无名的巴比伦人也许从未意识到，就算是在最荒唐的梦中也没有想到过，他的遗产在三千年后会变成什么样子。没人能够料到方程可解性的问题竟然催生了一个数学的核心概念，这就是群，或者说被证明为对称语言的群。更没有人会想到，对称会揭开物质世界的秘密。

解二次方程对物质世界的作用非常有限。只要这种解法只能用数字表示，而不能用符号（除非为了解方程而专门发明一种符号，而这样做只不过是为这一问题找一块遮羞布）表示，解五次方程就更没有多大用处。但是了解了五次方程为什么无法求解，理解了对称的关键作用，并将其潜在意义推而广之，就会打开通往整个物质王国之门。

对称理论的故事还将继续下去。对称理论对物理学的意义，事实上对整个科学的意义，仍然有待人们去发掘。我们还有很多东西需要了解，但是我们已经了解，至少在我们发现更强大的概念之前，对称群将仍是我们穿越未知世界的途径。

在物理学中，美自身并不能保证真的实现，而是有助于其实现。

而在数学中，美必须是真，因为只要是错误的，都是丑的。

参考读物

John C. Baez, "The octonions,"*Bulletin of the American Mathematical Society* volume 39 (2002) 145—205.

E. T. Bell, *Men of Mathematics* (2 volumes), Pelican, Harmondsworth, 1953.

R. Bourgne and J. -P. Azra, *Écrits et Mémoires Mathématiques d'Évariste Galois*, Gauthier-Villars, Paris, 1962.

Carl B.Boyer, *A History of Mathematics*, Wiley, New York, 1968.

W.K. Bühler, *Gauss: A Biographical Study*, Springer, Berlin, 1981.

Jerome Cardan, *The Book of My Life* (translated by Jean Stoner), Dent, London, 1931.

Girolamo Cardano, *The Great Art or the Rules of Algebra* (translated T.Richard Witmer), MIT Press, Cambridge, MA, 1968.

A.J.Coleman, "The greatest mathematical paper of all time,"*The Mathematical Intelligencer*, volume 11 (1989) 29—38.

Julian Lowell Coolidge, *The Mathematics of Great Amateurs*, Dover, New York, 1963.

P. C. W. Davies and J. Brown, *Superstrings*, Cambridge University Press,

Cambridge, 1988.

Underwood Dudley, *A Budget of Trisections*, Springer, New York, 1987.

Alexandre Dumas, *Mes Mémoires* (volume 4), Gallimard, Paris, 1967.

Euclid, *The Thirteen Books of Euclid's Elements* (translated by Sir Thomas L.Heath), Dover, New York, 1956 (3 volumes).

Carl Friedrich Gauss, *Disquisitiones Arithmeticae* (translated by Arthur A. Clarke), Yale University Press, New Haven, 1966.

Jan Gullberg, *Mathematics: From the Birth of Numbers*, Norton, New York, 1997.

George Gheverghese Joseph, *The Crest of the Peacock*, Penguin, London, 2000.

Brian Greene, *The Elegant Universe*, Norton, New York, 1999.

Michio Kaku, *Hyperspace*, Oxford University Press, Oxford, 1994.

Morris Kline, *Mathematical Thought from Ancient to Modern Times*, Oxford University Press, Oxford, 1972.

Helge S. Kragh, *Dirac—A Scientific Biography*, Cambridge University Press, Cambridge, 1990.

Mario Livio, *The Equation That Couldn' t Be Solved,* Simon & Schuster, New York, 2005.

J.-P. Luminet, *Black Holes*, Cambridge University Press, Cambridge, 1992.

Oystein Ore, *Niels Henrik Abel: Mathematician Extraordinary,* University of Minnesota Press, Minneapolis, 1957.

Abraham Pais, *Subtle Is the Lord: The Science and the Life of Albert Einstein,* Oxford University Press, Oxford, 1982.

Roger Penrose, *The Roat to Reality,* BCA, London, 2004.

Lisa Randall, *Warped Passages,* Allen Lane, London, 2005.

Michael I. Rosen, "Niels Hendrik Abel and equations of the fifth degree," *American Mathematical Monthly* volume 102 (1995) 495—505.

Tony Rothman, "The short life of Évariste Galois," *Scientific American* (April 1982) 112−120. Collected in Tony Rothman, *A Physicist on Madison Avenue,* Princeton University Press, 1991.

H. F. W. Saggs, *Everyday Life in Babylonia and Assyria*, Putnam; New York, 1965.

Lee Smolin, *Three Roads to Quantum Gravity*, Basic Books, New York, 2000.

Paul J. Steinhardt and Neil Turok, "Why the cosmological constant is small and positive," *Science* volume 312 (2006) 1180−1183.

Ian Stewart, *Galois Theory* (3rd edition), Chapman and Hall/CRC Press, Boca Raton 2004.

Jean-Pierre Tignol, *Galois's Theory of Algebraic Equations,* Longman, London, 1980.

Edward Witten, "Magic, mystery, and matrix," *Notices of the American Mathematical Society* volume 45 (1998) 1124−1129.

网站

A. Hulpke, Determining the Galois group of a rational polynomial:http://www.math.colosate.edu/hulpke/talks/galoistalk.pdf

The MacTutor History of Mathematics archive: http://www-history.mcs.st-andrews.ac.uk/index.html

A. Rothman, Genius and biographers: the fictionalization of Évariste Galois: http://godel.ph.utexas.edu/tonyr/galois.htm